# Spatially Regularized Reconstruction of Fibre Orientation Distributions in the Presence of Isotropic Diffusion

by

Quan Zhou

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2014

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

**Abstract**

The connectivity and structural integrity of the white matter of the brain is known to be implicated in a wide range of brain-related diseases and injuries. However, it is only since the advent of diffusion magnetic resonance imaging (dMRI) that researchers have been able to probe the miscrostructure of white matter in vivo.

Presently, among a range of methods of dMRI, high angular resolution diffusion imaging (HARDI) is known to excel in its ability to provide reliable information about the local orientations of neural fasciculi (aka fibre tracts). It preserves the high angular resolution property of diffusion spectrum imaging (DSI) but requires less measurements. Meanwhile, as opposed to the more traditional diffusion tensor imaging (DTI), HARDI is capable of distinguishing the orientations of multiple fibres passing through a given spatial voxel.

Unfortunately, the ability of HARDI to discriminate neural fibres that cross each other at acute angles is always limited. The limitation becomes the motivation to develop numerous post-processing tools, aiming at the improvement of the angular resolution of HARDI. Among such methods, spherical deconvolution (SD) is the one which attracts the most attentions. Due to its ill-posed nature, however, standard SD relies on a number of a priori assumptions needed to render its results unique and stable.

In the present thesis, we introduce a novel approach to the problem of non-blind SD of HARDI signals, which does not only consider the existence of anisotropic diffusion component of HARDI signal but also explicitly take the isotropic diffusion component into account. As a result of that, in addition to reconstruction of fODFs, our algorithm can also yield a useful estimation of its related IDM, which quantifies a relative contribution of the isotropic diffusion component as well as its spatial pattern. Moreover, one of the principal contributions is to demonstrate the effectiveness of exploiting different prior models for regularization of the spatial-domain behaviours of the reconstructed fODFs and IDMs. Specifically, the fibre continuity model has been used to force the local maxima of the fODFs to vary consistently throughout the brain, whereas the bounded variation model has helped us to achieve piecewise smooth reconstruction of the IDMs. The proposed algorithm is formulated as a convex minimization problem, which admits a unique and stable minimizer. Moreover, using ADMM, we have been able to find the optimal solution via a sequence of simpler optimization problems, which are both computationally efficient and amenable to parallel computations. In a series of both in silico and in vivo experiments, we demonstrate how the proposed solution can be used to successfully overcome the effect of partial voluming, while preserving the spatial coherency of cerebral diffusion at moderate to severe noise levels. The performance of the proposed method is compared with that of

several available alternatives, with the comparative results clearly supporting the viability and usefulness of our approach. Moreover, the results illustrate the power of applied spatial regularization terms.

**Acknowledgements**

To begin with, I would like to express my sincere gratitude towards my supervisor, Dr. Oleg Michailovich, for his continuous support and insightful advice throughout my entire research work and thesis writing. I truly appreciate for his expertise in identifying the problems and formulating the solutions in dMRI, patience in teaching, and attention in providing me with timely feedbacks. Without his guidance and encouragement, I would not have come even close to what I have achieved today.

My special thanks go to Iván Camilo Salgado Patarroyo and Eric Wai Yiu Lum for their invaluable suggestions and help with my research. It is a great honour to study and work with such talented companions.

Finally I would like to thank the researchers in the filed of dMRI. They have done great jobs which become the solid foundation of my research.

## Dedication

***This thesis is dedicated to my mother and wife.***

# Table of Contents

# List of Figures

# Chapter 1

# Introduction

## 1.1 Medical Imaging

"Natural science is the search for 'truth' about the natural world" [1]. As a branch of natural science, the objective of human medicine is to understand a particular object, the human body. Limited by our physical capabilities, our depth of understanding of human body partially depends on the tools which can expand the scope of our senses. For example, the foundation of anatomy was established on dissection of corpses owing to a lack of proper "ways" to look inside. But now, thanks to a variety of advanced imaging techniques, we can not only see in vivo images of the human body, but also can explore it on different levels, namely the atomic, molecular, cellular, tissue and whole-body levels.

Although the importance of medical imaging can not be neglected now, the original versions of those techniques were not designed for clinical medicine. Historically, medical imaging mainly derived from the defence and military areas once people recognized its potential uses in detecting and diagnosing human diseases and injuries. Now, the trend of medical imaging is shifting from a "technology push" pattern to a "biological/clinical pull" pattern. Doctors and researchers now are emphasizing the creating, improving and interpreting of different kinds of medical images.

Medical images generated by different techniques reveal a wide range of information about the human body. For example, X-rays yield mass density, ultrasound can give the blood volume and flow, and magnetic resonance imaging (MRI) senses the presence and density of water molecules.

## 1.2 Brain Imaging

Human brain, which is believed to be most "mysterious" part of human body, allows several imaging techniques to explore its secrets. X-ray computed tomography (CT) is one of those available techniques, which is comprised of a series of x-ray beams passing through the head. With the help of CT, we can view the structure of brain but not its function. Moreover its x-ray related nature implies that CT shares the same radiation risks as x-ray. An alternative method, positron emission tomography (PET), can provide the functions of different regions of brain, but replies on radioactive material which need to be injected or inhaled to blood prior to imaging procedure. MRI does not reply on injected external material or radioactive rays. It senses the radio frequency signals emitted from "triggered" water molecules, specifically hydrogen atoms. It provides different ways to capture both the structure (diffusion MRI) and function (functional MRI) of human brain. For example, the up-to-date achievement of functional MRI is that human thoughts can be visualized on computer screen by analysing and interpreting the results of its functional MRI. The major interest of this thesis is focused on diffusion MRI, one branch of MRI, which can delineate the microstructure of brain.

## 1.3 MRI

Water, occupying 65% of human body on average, is hypothetically believed to be the major reason of the existence of life. The understanding of its importance stayed on the level of supporting life until the discovery of nuclear magnetic resonance (NMR) which allowed detecting and quantifying the presence and density of water in human body. In the last few decades people never stop their pace to chase the secrets of life on a microscopic level with the help of NMR.

Magnetic resonance imaging as a mature application of NMR technique provides a noninvasive way to sense the presence and properties of water in human tissues. Those information can act as sensitive indicators which play a vital role in diagnosing diseases and injuries. But the information water can give us is not limited within that. In middle of 1980s, the random motion of water, aka Brownian motion, was "captured" and measured by a novel MRI technique, diffusion MRI (dMRI). Since then, dMRI has encompassed a number of various methodologies and protocols which aim to reveal the microstructure of human brain.

## 1.4 dMRI

The connectivity and structural integrity of white matter in human brain is nowadays known to be indicative of a wide range of brain-related pathologies. While "invisible" to alternative means of imaging based diagnosis, the above information can be elicited from the measurements acquired by means of dMRI. This fact has triggered an active development of various dMRI methodologies, which have made dMRI into a well-established technique of modern medical imaging [2].

At the present time, dMRI encompasses a number of various methodologies and protocols, the most widely acknowledged of which is diffusion tensor imaging (DTI) [3, 4, 5, 6, 7]. However, the capacity of DTI is limited due to its reliance on assuming the ensemble averaged diffusion propagator (EAP) to be a unimodal Gaussian. In fact, the above assumption undermines the ability of DTI to provide accurate estimation of the apparent diffusivity of white matter at the locations of crossing, diverging, and kissing neural fibre tracts [8, 9, 10]. Alternatively, a parameter-free approach to estimation of the EAP is provided by diffusion spectral imaging (DSI) [11, 12], which allows reconstruction of complex diffusivity profiles under rather general conditions. Unfortunately, practical implementation of DSI entails acquisition of diffusion measurements over a dense Cartesian grid in the $q$-space (defined in Section 2.4), which renders the acquisition requirements of DSI beyond the limits of practically admissible. However, this problem can be alleviated by restricting the diffusion measurements to a relatively small number of concentric shells in the $q$-space. This sampling strategy – known as multi-shell diffusion imaging (MSDI) [13, 14, 15] – has successfully served as a basis for many advanced dMRI methodologies [16, 17, 13, 18, 19, 20, 21, 22].

## 1.5 HARDI

Although the availability of the EAP is generally preferred, in some cases it is sufficient to know the result of its marginalization over the range variable. The resulting probability density is known as the orientation distribution function (ODF). It quantifies the probability with which water molecules undergo displacement along various spatial directions [23, 24]. On the practical side, a useful approximation of the ODF can be obtained by means of $q$-ball imaging (QBI) [24], which can in turn be based on a single-shell data acquisition scheme, known as high angular resolution diffusion imaging (HARDI) [25, 26, 27, 28, 9, 29, 30]. The adoption of Funk-Radon transform (FRT) in QBI [24] introduces a side effect, zeroth-order Bessel function, which imposes constraints on the angular resolution of estimated ODFs, especially for applications like multi-fibre tractography [31].

The problem of limited angular resolution of QBI can be addressed using the framework of spherical deconvolution (SD) [32]. In this formulation, SD is used to recover a fibre orientation distribution function (fODF) which, as opposed to an ODF, quantifies the likelihood of a neural fibre to have a (local) tangent vector of a certain orientation. The idea of using SD as a method for improving the angular resolution of QBI was first introduced in [33] through a direct matrix inversion preceded by low-pass filtering. A more advanced SD technique based on a least-square (LS) formulation with non-negativity constraint was proposed by the same authors in [34, 35]. In [36], the fODF is parametrized based on a two-compartment model, followed by its reconstruction using non-regularized (naive) inverse filtering. A similar line of arguments is used in [37], albeit this time with Tikhonov regularization implemented implicitly through damped SVD-based inversion. Advanced statistical considerations were also employed in [38] to derive a maximum entropy deconvolution algorithm. The same idea of entropy maximization was later adopted by [39]. In the latter work, however, the reconstruction was performed within the framework of blind source separation, which offers the substantial advantage of independence on the knowledge of single fibre response (SFR). Unfortunately, the approach used by the authors in [39] does not explicitly take into account the distribution and level of measurement noises, which makes it overly dependent on measurement conditions. Finally, we would also like to mention the SD approach of [40], which is close in philosophy to the methods discussed in this thesis. In particular, [40] advocates the idea of sparse SD, congruent with the assumption on an fODF to have a relatively small number of "sharp" maxima in the directions of associated neural bundles.

Obviously, the above-mentioned methods intentionally neglect the effect of isotropic diffusion. This deficiency has been alleviated in a series of more recent works on the subject. Thus, in [41, 42], fODF has been supplanted by a different quantity, called a fibre orientation function (FOF), which explains a combined effect of both (multimodel) anisotropic and isotropic diffusions. The deconvolution algorithm in [43] iteratively estimates and subtracts the isotropic component from HARDI data before fODFs are estimated. A data formation model similar to that of [41, 42] was also explored by the authors of [44]. Quite a different type of deconvolution methods were explored in [45, 46] based on Bayesian inference. Specifically, [46] takes advantage of a measure-theoretic framework to represent the fODF as a probability measure decomposable into three unique components, which allow adequate description of a wide spectrum of possible fODF shapes and patterns. Similarly to the previously mentioned results, the models used in [45, 46] explicitly account for the presence of isotropic diffusion.

Since neutral bundles extend continuously in space, it is reasonable to assume spatially adjacent fODFs to exhibit a fair amount of correlation, which can be exploited to improve

the results of SD. This idea has been exploited in several studies to improve the accuracy of QBI [47, 48]. In application to SD, spatial regularization was used in [49], in which case a spatial continuity of fODFs was enforced via minimizing a weighted quadratic penalty. A more recent work in [50] has introduced an anisotropic regularization scheme which, for any spatial direction $\mathbf{u}$, minimizes the $\mathcal{L}_2$-norm of the projection of the spatial gradient of fODFs onto $\mathbf{u}$. It has been shown that this minimization favours the spatial continuity of neural fibres, thereby producing anatomically plausible reconstructions. Unfortunately, neither of the above-mentioned approaches was designed to deconvolve HARDI signals in the presence of isotropic diffusion.

## 1.6 Contributions of the Thesis

Despite the apparent success of SD in application to diffusion imaging, there is still much space for further improvements. In particular, it appears that little has been done on the development of SD algorithms which can perform reliably in the presence of isotropic diffusion, while imposing effective regularization constraints on both isotropic and anisotropic components of diffusion signals. Contriving such a reconstruction problem is likely to result in composite optimization, solving which could be a non-trivial problem by itself. Accordingly, the research of this thesis aims to contribute to the existing body of works on SD in the following directions. We formulate a new method for SD of HARDI data, subject to spatial regularization of both isotropic and anisotropic components of HARDI signals as well as their related fODFs. We describe a computationally efficient implementation of the proposed algorithm based on the idea of variable splitting [51]. The proposed computational solution has a particularly simple modular structure, which is straightforward to reproduce using standard computational means.

## 1.7 Organization of the Thesis

The remainder of the thesis is organized as follows. Chapter 2 starts with the basic principles of MRI and extends it to the field of dMRI. Then chapter 3 specifically focuses on the existing methods which deal with a particular type of dMRI, HARDI, which is the core interest of this thesis. Chapter 4 details the data formation model, methodology, and analytic solution of proposed method, spatially constrained sparse deconvolution (SCSD). Chapter 5 sets out the experimental setup used for its numerical validation. The results of both computer simulated and in vivo data are reported in it as well. The chapter also

contains a comparative analysis of the performance of the proposed algorithm against that of several alternative solutions. Finally, Chapter 6 summarises the main contribution of this thesis and Chapter 7 leads the discussion to future research directions.

# Chapter 2

# MRI and dMRI

In 1946, F. Block and E. Purcell discovered the phenomenon of nuclear magnetic resonance (NMR), which nowadays is a fundamental technique in medicine care [2, 52, 53, 54, 55, 7, 56]. Based on what they had found, a Nobel Prize was issued in 1952. The NMR extended its glory with a new medical imaging technique, nuclear magnetic resonance imaging (NMRI) which was invented in 1970s and then became a standard imaging technique for both academic and clinical usages. To relieve public concern and keep distance from nuclear energy, magnetic resonance imaging (MRI) was kept as the official name instead of NMRI.

For the purpose of clinical medicine care, the traditional MRI is implemented on different tissues of patients to help the doctor diagnose. Despite its success in the past decades, the relatively "simple" two-dimensional images are insensitive to microstructure of neural fibres of the human brain which limits its further usage [2, 7]. The limitation became the direct motivation to improve and extend the function of MRI. Bihan is believed to be the first one who proposed to use MRI to sense the water diffusion in [6] which later was proved to be a milestone. Diffusion MRI allows one to infer the geometrical structure of neural fibres in the white matter of human brain which can indicate brain-related disorders [7].

This chapter, we go through the basic concepts of MRI and it derivation, dMRI. Meanwhile special attention will be paid on different types of dMRI signal sampling schemes, namely DTI, DSI, and HARDI.

## 2.1 Principles of MRI

Water, the major component of human body, makes up 70% to 90% of most tissues [57]. Therefore the presence and density of water can work as an indicator of diseases or abnormal changes of tissues [57, 2]. Thanks to the discover of NMR, we can excite the water molecule, to be specific hydrogen nuclei, to spin with respect to a direction by applying or changing the strength of magnetic field. The generation of molecule spins and their changes can be measured as signals which will be collected for further process or analysis. Taking the strengths of signals as the intensities of pixels in corresponding spatial locations, we can "draw" a image of NMR signals which is the imaging procedure of MRI results. In this section, the brief explanation will be presented about the mechanism of MRI.

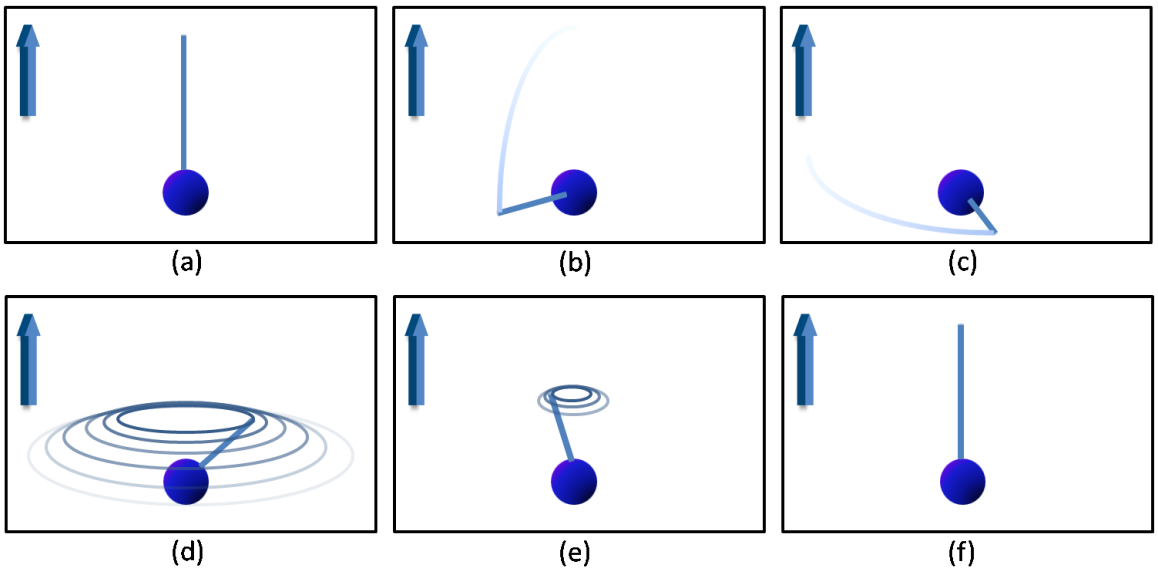


Figure 2.1: Status of Hydrogen nucleus spin with different magnetic fields. (a) spin aligned with static magnetic field $\mathbf{B}_0$ indicated by white upright arrow, (b) transaction after application of $\pi/2$ RF pulse, (c) transverse spin right after $\pi/2$ RF pulse, (d) transverse projection of relaxation, (e) longitude projection of relaxation, and (f) the end of relaxation, ready for next RF pulse. This is a redrawing of Fig. 2.1 in [2, p.13].

Hydrogen nucleus, of particular importance in MRI for its abundance of existence in human tissue, possesses a magnetic dipole which is normally referred as a spin [2, 53, 52, 55, 56, 54]. When put in a strong static external magnetic field $\mathbf{B}_0$ generated by MRI scanner, these magnetic spins will align themselves with respect to the direction

of external static magnetic field and precess, shown in Fig. 2.1 (a). Not only the north and south pole directions change accordingly, the frequencies of precession of spins also keep a linear relationship with the strength of static magnetic field. The proportional relationship between $\mathbf{B}_0$ and precession radian frequency $\omega$ can be expressed by Larmor equation [56, 7, 2],

$$\boldsymbol{\omega} = \gamma \mathbf{B}_0, \tag{2.1}$$

where $\boldsymbol{\omega}$ is formally referred to Larmor or resonant frequency and $\gamma = 2.675 \times 10^8$ radians/tesla is known as gyromagnetic ratio, a constant value for hydrogen nucleus, $H^1$.

Applying only one external magnetic is not enough to excite nuclei to generate measurable signals because the magnetic moment vectors of hydrogen nuclei have no preference orientation. Therefore a second type of magnetic field will be generated for a brief duration and it will oscillates at radio frequencies after which it gets the name, RF pulse. The usages of RF pulses have two main purposes: excitation and refocusing. To distinguish the usages and functions of RF pulses, they are named with the spin orientation changes they cause. To be specific, a RF pulse is named a $\theta$ RF pulse if it changes the orientation of spins by $\theta$ radian, e.g. $\pi/2$ RF pulse.

When applying a $\pi/2$ RF pulse on the perpendicular plane of $\mathbf{B}_0$ with the Larmor frequency, the spins will move away from $\mathbf{B}_0$ to precess on the perpendicular plane, illustrated in Fig. 2.1 (b)(c). After $\pi/2$ RF pulse, the projection of spins on perpendicular plane, Fig. 2.1 (d), will decay exponentially with respect to a constant time $T_2$ whose is counted in 100 milliseconds. Assuming in Cartesian coordinates a static magnetic field is applied in the direction of $z$ and denoting the time-changing magnetization vector generated by spins with $\mathbf{M}$, the exponential relationship mentioned here can be written as [56, 55],

$$\frac{d\mathbf{M}_{xy}}{dt} = -\frac{\mathbf{M}_{xy}}{T_2} \quad \text{or} \quad \mathbf{M}_{xy} = \mathbf{M}_0 e^{-t/T_2}, \tag{2.2}$$

where $\mathbf{M}_{xy}$ is the projection of $\mathbf{M}$ on $xy$ plane and $\mathbf{M}_0$ is the initial value of $\mathbf{M}_z$, the projection of $\mathbf{M}$ on $z$ axis. Analogously, the recovery time of spins on $\mathbf{B}_0$ direction has a exponential expression with respect to another time constant $T_1$ who is typically on the order of 1 second. With the same assumption as previous one, the relationship can be written as [56, 55],

$$\frac{d\mathbf{M}_z}{dt} = \frac{\mathbf{M}_0 - \mathbf{M}_z}{T_1} \quad \text{or} \quad \mathbf{M}_z = \mathbf{M}_0(1 - e^{-t/T_1}). \tag{2.3}$$

$T_1$ and $T_2$ together decide the relaxation time of MRI which are also the basis of $T_1$ and $T_2$ weighted MRI contrasts.

So far we have learned that we can excite the water molecules to spin according to the applied external magnetic field. The next step is to sense or reflect those excited spins. According to Faraday's law, the rotating magnetic field generated by coherently precessing spins can induce a current in the receiver coil in MRI scanner. Moreover the strength of current can be reflected as pixel intensity in a grey image which is the visible result of MRI. It is important to notice that the brightness of each single pixel is decided by the coherence of the phases of spins. Due to the inhomogeneities of the first $\pi/2$ RF pulse, the frequencies of spins vary slightly. When the time passes, the coherence of spin phases and induced current signal decay gradually. The loss of signal and coherence of phases is called $T_2$ signal loss. To strengthen the signal and realign the spins, a second flipping pulse, $\pi$ pulse, is applied at time $t$ after the first excitation $\pi/2$ pulse. The purpose of second RF pulse is to flip every spin by $\pi$ on the perpendicular plane of $\mathbf{B}_0$, in which case the faster spins will be put behind the slower spins. And after another $t$ time, all spins keep the same pace with each other again, which is called refocusing. This moment, when we can observe spin coherence again, is called a spin echo.

The third group of magnetic fields are magnetic field gradients or simply gradients. Like the RF pulse, gradients are also applied during short periods in the whole acquisition process. Normally, in Cartesian coordinates, the gradients can be represented by $G_x(t)$, $G_y(t)$, and $G_z(t)$ which create linear magnetic field changes along three orthogonal directions. Taking the gradient components into account, the resulted magnetic field can be written as,

$$\mathbf{B} = \mathbf{B}_0 + G_x(t)\mathbf{x} + G_y(t)\mathbf{y} + G_z(t)\mathbf{z}, \tag{2.4}$$

where $\mathbf{x}, \mathbf{y}, \mathbf{z}$ stand for unit vectors pointing three orthogonal directions. The application of gradients can achieve three goals [2]: (1) slice selection, (2) image encoding, and (3) diffusion weighting (discussed in the next section).

Let us talk about the first function of gradient : slice selection. Slice selection is done by applying RF pulses during the simultaneous application of a gradient. According to (2.1), Larmor frequencies of spins change linearly on the gradient direction and only those spins who are located on the same plane (slice) perpendicular to gradient direction will possess the same Larmor frequency. The simultaneous RF pulses act as frequency selectors, exciting and refocusing only one slice on which the spins have the Larmor frequency same as the RF pulses. The RF pulses and slice selection gradient work together to reduce the three-dimensional volume into a two-dimensional slice.The thickness of slice is decided by the frequency bandwidth of RF pulses and the strength of the gradient [2].

For example, on top of static magnetic field $\mathbf{B}_0$, a linear gradient is applied pointing to north pole, the direction of $\mathbf{z}$ in Cartesian coordinates, with simultaneous $\pi/2$ RF pulses.

Only the spins on one $xy$ plane (slice) will be excited because the Larmor frequencies of those spins are the same as the applied $\pi/2$ RF pulse. Up to this point, away from visible two-dimensional MRI image, there is one more step to go: to distinguish each pixel on the excited slice. Actually it can be done by applying linear gradients ($G_x$ and $G_y$) along two orthogonal in-plane directions ($\mathbf{x}$ and $\mathbf{y}$) after the spins are excited. The linear gradient changes cause the Larmor frequencies change according to the pixel location, $(x, y)$. In another word, each pixel on the excited slice is labelled by a distinct Larmor frequency. With time passing, the accumulated phase $\theta$ of spin at $(x, y)$ on the slice can be expressed by [2],

$$\theta = 2\pi(k_x x + k_y y), \tag{2.5}$$

where $k_x = \gamma \int G_x(t)dt$ and $k_y = \gamma \int G_y(t)dt$. It implies that the net magnetization on the selected slice can be formulated as a function of pixel locations, e.g. $f(x, y)$. Therefore the overall integrated current signal sensed by receiver coil can be written as [2],

$$s(t) = \int\int f(x, y)e^{i2\pi(k_x x + k_y y)}dxdy, \tag{2.6}$$

which is also know as Fourier transform. It illustrates the Fourier relationship between $s(t)$ and $f(x, y)$. Once $k_x$ and $k_y$ are determined by MRI scanner during the acquisition process, we can recover $f(x, y)$ by inverse Fourier transform of $s(t)$. Since the signal $s(t)$ is measured according to the values of $k_x$ and $k_y$, based on which we can define a "K-Space" to represent different values of $k_x$ and $k_y$. The choice of $k_x$ and $k_y$ on "K-Space" defines measurement path or trajectory. Moreover it decides the quality of result images.

## 2.2 Principles of dMRI

The definition of diffusion in [2] is: "Diffusion is a mass transport process arising in nature, which results in molecular or particle mixing without requiring bulk motion." Before we step into details of dMRI , some important points of diffusion in MRI need to be clarified first. First of all, diffusion describes the random motion of small articles, e.g. molecules, instead of bulk motion from one place to another. Secondly, we are interested in the averaged diffusion motion not focusing on a specific molecule. Finally, the water diffusion phenomenon in brain can not be directly measured. As a matter of fact, the signal loss caused by water diffusion within human tissues is the measurable phenomenon.

To quantify the signal loss caused by water molecule diffusion, a group of linear varying gradients need to be applied in order to dephase and rephase spins. A pair of dephasing

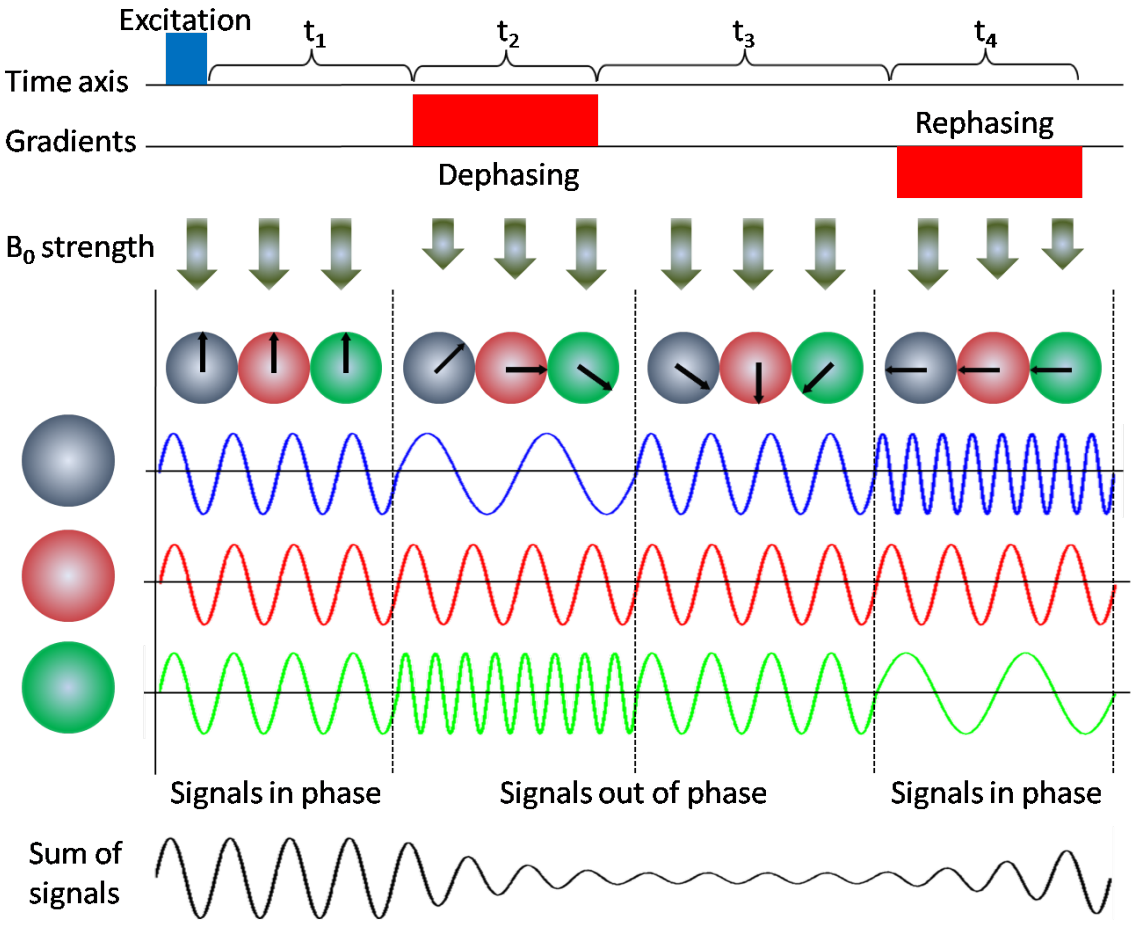


Figure 2.2: An example of the application of dephasing and rephasing gradients after excitation RF pulse. Red, green, and blue circles represent different water molecules at adjacent locations. Black thick arrows indicate different strengths of magnetic fields. Black thin arrows give the phases of MRI signals from different regions of molecules. This is a redrawing of Fig 1.9 in [7, p.9].

and rephasing gradients are shown in Fig. 2.2. During time period $t_1$ which is right after applying exciting RF pulse, three regions of spins possess the same Larmor frequency with the coherent phase. When stepping into $t_2$ period, three regions of spins change their Larmor frequencies according to the change of magnetic fields. The changes of frequencies then are reflected by the accumulated phases during $t_3$. The second gradient acts as the opposite of the first one. It accelerates the slower spins and slow down the faster ones to

accomplish the rephasing mission, which aims to realign the phases of spins at different regions at the end of time $t_4$. At the first glance, it seems that this pair of dephasing and rephasing gradients barely change anything during the whole process. As a matter of fact, the signal we can measure at the end of $t_4$ is diffusion weighted. This is because, the perfect rephasing can only be achieved if there is no diffusion motion of water molecules between the application of two gradients. Unfortunately water molecules always diffuse which will "contaminate" the realigned phases. In theory, the imperfect phases can be detected individually for they are different from others. But in practise, we quantify them by measuring the signal strength loss instead. Following the gradient application procedure, the resulted MRI images are called diffusion weighted images because the pixel intensities reflect the water molecule diffusion. And those applied gradients are named diffusion weighting gradients. Specifically, the pair of dephasing and rephasing gradients used in Fig. 2.2 are called bipolar gradients.

Though the application of bipolar gradients allows water diffusion to show itself by decaying measured signals, it has drawbacks. One of them is that the signal loss is not only related to water molecule diffusion but partially owing to $T_2$ decay.

To eliminate the effect of $T_2$ decay, a $\pi$ RF pulse is used after the dephasing gradient. Since the additional RF pulse changes the sign of spin phases, the rephasing gradient needs to have the same sign as the first one, in which case, they both together are called unipolar gradients. Fig. 2.3 shows an example of unipolar gradients along with same parameters which affect the signal loss in dMRI. This arrange of RF pulses and diffusion gradients is formally called pulsed gradient spin-echo (PGSE) which was proposed by Stejskal and Tanner (1965) [2, 58]. We can use $S_0$ to denote the MRI signal acquired without diffusion

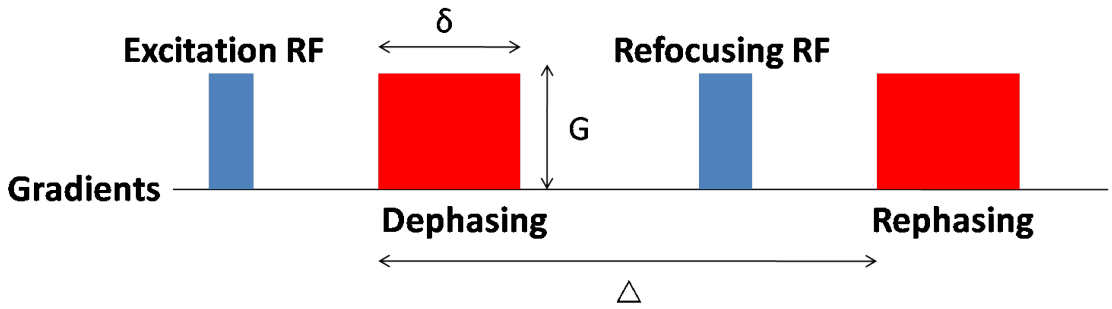


Figure 2.3: Example of unipolar gradients. $\delta$ is the the duration of the gradient pulses, $\Delta$ is the the mixing time, and G is the strength of the gradients.

weighting which will be later used to do normalization. And the one to be normalized is

the diffusion weighted signal $S$ whose strength is determined by several parameters defined in Fig. 2.3. Intuitively, the longer mixing time $\Delta$ we use, the more misaligned phases will be, which indicates weaker signals. Moreover, $\delta$ and $G$ together decide the amount of initial dephasing. The forth parameter, the apparent diffusion coefficient (ADC) $D$, is not controlled by scanner but the nature of water molecule diffusion. Putting all together, the signal attenuation can be expressed as a function of four parameters: $S/S_0 = f(\Delta, D, G, \delta)$ [7]. Manipulating the values of three parameters $(\Delta, G, \delta)$ generated by MRI scanner can achieve different degrees of diffusion weighting. No matter what choices are made, the final purpose aims to recover the apparent diffusion coefficient $(D)$, which reveals the nature of diffusion motion along the direction of diffusion gradient. Analysing the effects of controlled three parameters on spin phases, the signal attenuation can be expressed accordingly as,

$$\frac{S}{S_0} = e^{-\gamma^2 G^2 \delta^2 (\Delta - \delta/3) D}, \tag{2.7}$$

with $\gamma$ being gyromagnetic ratio as in (2.1). From the perspective of signal processing, scanner controlled parameters can be regarded as constant values predefined. Therefore we can here introduce a new parameter, $b$-value, to represent them all, $b = \gamma^2 G^2 \delta^2 (\Delta - \delta/3)$. So again, the signal loss can be simplified as,

$$\frac{S}{S_0} = e^{-bD}. \tag{2.8}$$

Notice that the ADC, we can recover from measurements of $S$ and $S_0$ by (2.8), only represents the characteristics of water molecule diffusion on one direction, aligned with diffusion gradient. To reveal a complete representation of water molecule diffusion, more complicated models and more measurements with different gradient directions will be needed, which will be discussed in following sections.

## 2.3 DTI

Supposing the water molecules can diffuse freely, then the ADCs recovered would be isotropic. In another word, the ADCs have no preference of orientations and they share the same value for every orientation. We can define that kind of diffusion as isotropic diffusion against the other pattern, anisotropic diffusion. When obstacles exist on the path of water diffusion, the movement will be blocked which will be reflected in the corresponding recovered ADCs. In that case, diffusion has its own directionality which explains the name of that diffusion pattern, anisotropic diffusion. The anisotropic diffusion is of great interests

for it carries the indirect information of living tissue microstrcture where water molecule diffusion happen. To be specific, water molecules "prefer" to diffuse along organized tissues or protein filaments. This preference would be revealed by the corresponding ADC from which we can deduce the existence of such structure.

According to the explanation of diffusion weighted MRI signal acquisition procedure, we need to measure an $S$ for each orientation and one extra $S_0$ at least in order to calculate the corresponding ADC. Thanks to its simplicity, isotropic diffusion can be represented with only one ADC which asks for one $S$ and one $S_0$. If we need a complete representation of anisotropic diffusion, in theory, infinite number of measurements are required. That is not only inefficient but also not economic because it ignores the similarities between ADCs with adjacent orientations.

The first trial to model anisotropic diffusion is DTI which was firstly introduced by Basser in his papers [3, 4] in 1994. The success of DTI relies on the fact that, for the first time, anisotropic diffusion can be extracted, characterized, and exploited. DTI provides an opportunity to probe the microstructure of neural fibres. Furthermore, DTI expanded it influence on other methods dealing with more complex diffusion signals. For example, in a number of spherical deconvolution methods, DTI was adopted as the model for single fibre response.

DTI, as its name implies, relies on diffusion tensor which can characterize Gaussian diffusion. The diffusion tensor is mathematically expressed by a $3 \times 3$ symmetric matrix,

$$\mathbf{D} = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{bmatrix}. \tag{2.9}$$

And accordingly, the diffusion signal attenuation can be written as,

$$\frac{S}{S_0} = e^{-b\mathbf{u}^T\mathbf{D}\mathbf{u}}, \tag{2.10}$$

where $\mathbf{u} \in \mathbb{S}^2$ is an arbitrary unit vector indicating direction. Take a close look at the elements of $\mathbf{D}$. The elements on diagonal directly stand for the diffusivities along three orthogonal axes in Cartesian coordinates. The off-diagonal elements correspond to the correlation between displacements along those orthogonal axes. It would be easy to understand the concept by regarding diffusion tensor as a three-dimensional covariance matrix of displacements during a given time period. Furthermore, the diffusion tensor $\mathbf{D}$ can fully describe the characteristics of an ellipsoid in Cartesian coordinates. And the ellipsoid is the contour of diffusivities if we draw ADCs as a function of corresponding gradient directions. Since there are only six unknown elements in $\mathbf{D}$, in theory, we only need to take

seven measurements (including $S_0$) to calculate $\mathbf{D}$. To make the estimate of $\mathbf{D}$ robust, in practise, more measurements are required as the input of lease square calculation [2].

The analysis and usage of resulting diffusion tensor mainly focus on the eigenvalues and eigenvectors of $\mathbf{D}$ which indicate the directions of anisotropy. That is not the end of story. There are two types of contrasts, generated from $\mathbf{D}$, are commonly used in grey images, averaged ADC and anisotropy map [7]. Averaged ADC can be obtained by averaging the eigenvalues of $\mathbf{D}$, e.g. $\bar{\lambda} = (\lambda_1 + \lambda_2 + \lambda_3)/3$. And the most popular representative of anisotropy map is fractional anisotropy (FA), which is defined as [7]:

$$FA = \sqrt{\frac{3}{2}}\sqrt{\frac{(\lambda_1 - \bar{\lambda})^2 + (\lambda_2 - \bar{\lambda})^2 + (\lambda_3 - \bar{\lambda})^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}. \tag{2.11}$$

Of course, FA is not the only map which illustrate the diffusion anisotropy. There are two more ways to show it [7]:

$$RA = \sqrt{\frac{1}{2}}\frac{\sqrt{(\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_3 - \lambda_1)^2}}{\lambda_1 + \lambda_2 + \lambda_3} \tag{2.12}$$

$$VR = \frac{\lambda_1\lambda_2\lambda_3}{(\bar{\lambda})^3} \tag{2.13}$$

where RA stands for the relative anisotropy and VR is the volume ratio.

Despite of its unparalleled success [5, 16, 7] and diversity of contrasts, DTI has a significant limitation which directly arise from the fundamental assumption, Gaussian diffusion model. Working along with this assumption, DTI can only perfectly reveal the nature of single neural fibre within one voxel [24, 33, 38, 42]. For voxels accommodating two or more fibres crossing, bending or twisting each other, DTI can not yield precise description. Therefore, researchers adopted more complex models or model-free methods to improves the resulted diffusion representation.

## 2.4 Q-Space Imaging and DSI

Before detailed discussion of $\mathbf{q}$-space imaging and DSI, several definitions need to be clarified. When defining $b$-value in section 2.2, we discuss that the normalized diffusion weighted signal depends on three scanner controlled parameters: $\Delta$ , $G$, and $\delta$. Besides $b$-value, we can also define $\mathbf{q}$-value or wave-vector, $\mathbf{q} \in \mathbb{R}^3$ [2],

$$\mathbf{q} = \frac{\gamma\delta G\mathbf{u}}{2\pi}, \tag{2.14}$$

which can help the determination of spin phase along with displacement distance $\mathbf{x}$. Since the value and orientation of $\mathbf{q}$ can be manipulated by changing the corresponding diffusion gradient, we can define a three-dimensional $\mathbf{q}$-space based on which the normalized diffusion signal, $E(\mathbf{q})$, is defined.

The second definition is diffusion propagator $P(\mathbf{x}_1, \mathbf{x}_2, \Delta)$ quantifying the likelihood that a water molecule randomly moves from predefined initial location $\mathbf{x}_1 \in \mathbb{R}^3$ to a new location $\mathbf{x}_2 \in \mathbb{R}^3$ within the time duration $\Delta$. As mentioned at the beginning of section 2.2, we are interested in the ensemble movement of water molecule diffusion not a particular one. Therefore we define ensemble average propagator (EAP), $\bar{P}(\mathbf{x}, \Delta)$ , based on $P(\mathbf{x}_1, \mathbf{x}_2, \Delta)$ [2],

$$\bar{P}(\mathbf{x}, \Delta) = \int_{\mathbb{R}^3} \rho(\mathbf{x}_1) P(\mathbf{x}_1, \mathbf{x}_2, \Delta) d\mathbf{x}_1. \tag{2.15}$$

where $\rho(\mathbf{x}_1)$ is initial spin density before the diffusion gradients are applied and $\mathbf{x} = \mathbf{x}_2 - \mathbf{x}_1$. With the definition of $\mathbf{q}$, we can use the term $e^{i2\pi\mathbf{q}\cdot\mathbf{x}}$ to express the phase change of spin when it moves from $\mathbf{x}_1$ to $\mathbf{x}_2$, where "$\cdot$" stands for inner product. Multiplying EAP to the phase change, we can obtain the averaged phase change $\bar{P}(\mathbf{x}, \Delta)e^{i2\pi\mathbf{q}\cdot\mathbf{x}}$. Then the normalized diffusion signal is obtained by [11, 2, 24]

$$E(\mathbf{q}) = \int_{\mathbb{R}^3} \bar{P}(\mathbf{x}, \Delta) e^{i2\pi\mathbf{q}\cdot\mathbf{x}} d\mathbf{x}, \tag{2.16}$$

as a function of $\mathbf{q}$. The above equation illustrates a Fourier relationship between $E(\mathbf{q})$ and $\bar{P}(\mathbf{x}, \Delta)$ from which we can tell that it is possible to recover $\bar{P}(\mathbf{x}, \Delta)$ through the inverse Fourier transform of normalized diffusion signals. The whole process was firstly proposed by Stejskal in [59] which is nowadays called $\mathbf{q}$-space imaging. It can reveal the microscopic structure of underlying neural fibres by reconstructed EAP. Moreover, the $\mathbf{q}$-space imaging allows for resolution of intravoxel fibres crossing and diverging each other. But it has limitations [11]: (1) The long measurement time required by enormous diffusion signals needs stable experimental subject which can not be achieved in vivo. The biological motion and eddy current distortions will contaminate the phase of the echo signal; (2) extreme high diffusion gradients are not available on clinical scanners.

Diffusion spectrum imaging (DSI) was briefly introduced in [60], then fully described in Tuch's dissertation [11] in 2002. DSI is related with $\mathbf{q}$-space imaging but differs in some aspects. According to the analysis in [11] and [2], DSI embeds the diffusion weighting gradients into the imaging protocol and EAP is reconstructed at each voxel of the image. DSI takes the measurements of diffusion signals on a grid in $\mathbf{q}$-space, shown in Fig. 2.4(a), which allows the usage of fast Fourier transform (FFT). The resulting EAP is also com-

prised of samples on the grid. Then the final objective result, orientation distribution function (ODF), is obtained by projecting EAP onto unit sphere.

Even it derives from **q**-space imaging and improves the performance, DSI still has some drawbacks [2]. The major limitation of DSI is the acquisition time. Compared with DTI (require minimum seven measurements), the typical number of DSI measurements is 500-1000. Furthermore, the resolution of resulting EAP and ODF depends on infinitely short pulses which are unrealistic. In practise, the side effect of short pulses applied in DSI is blurring results.

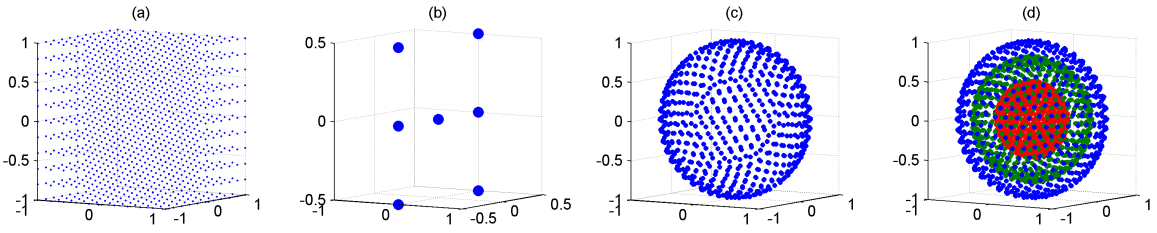


Figure 2.4: Examples of different acquisition schemes: (a)DSI, (b)DTI, (c) HARDI, and (d)Multishell.

## 2.5 HARDI

High angular resolution diffusion imaging (HARDI) was first proposed by Tuch in [61] as a new diffusion sampling scheme in **q**-space. Unlike DTI taking several measurements in **q**-space Fig. 2.4 (b) or DSI taking samples according to a grid of **q**-space Fig. 2.4 (a), HARDI was introduced as a strategy choosing 126 directions on a 5-fold tessellated icosahedral hemisphere in the papers [61, 11]. Therefore, compared with the other two, HARDI possesses a medium measurement requirement, typically 50-100 measurements, which is feasible for in vivo data. Compared with DSI, the results of HARDI, either ODF or fibre orientation distribution function (fODF) (detailed information in next chapter), is always spherical function, same as the sampled signals. Especially, resulting ODF is the radial projection of EAP. Although the results are reduced to be two-dimensional, they preserve high angular resolution property which can distinguish multiple fibres within one voxel. Meanwhile, HARDI is not designed based on model assumption like DTI, so it allows model-independent and model-independent methods both. The early trail with HARDI is to fit multi-compartment models [10, 16, 62, 63, 49]. Then a **q**-ball imaging (QBI) was

proposed in [24] which does not reply on any format of diffusion model at all. To make QBI robust to noise, improving versions of QBI showed up with additional regularizers [29, 27, 26, 64] (discussed in the next chapter). Moreover, some researchers attempted to utilize diffusion signals sampled on multiple shells in **q**-space to make use of different properties of each shell, Fig. 2.4 (d).

## 2.6 Conclusion

The discovery of NMR opens a gate for medical imaging researchers to accurately measure the presence and density of water molecules in human tissues. Relying on different relaxation mechanism, $T_1$ and $T_2$, two types of grey images can be obtained for different diagnosis purposes.

dMRI as a special branch of MRI draws more and more attentions for its unique ability to probe microstructure of white matter of human brain. On top of the fundamental excitation and refocusing RF pulses, additional diffusion gradients are inserted into the whole acquisition procedure to allow the MRI signal to sense the water molecule diffusion motion. The resulting signal is so called diffusion weighted, based on which ADC can be calculated accordingly.

A single ADC can not be used for higher level application. So the early trial DTI uses a group of diffusion weighted signals, sampled at different locations in **q**-space, to fit a Gaussian diffusion model in order to estimate related diffusion tensor. Not only the eigenvectors and eigenvalues can indicate the major diffusion orientation, but also a group of related measures (such as FA and RA) can provide further information about underlying neural fibres.

The major limitation of DTI encourages the creative thinking about more complicated methods to resolve multiple fibres within one voxel. The proposal of DSI theoretically provide a solution to that issue, which generates EAP by iFFT of diffusion signals. Although EAP is capable of distinguishing multiple fibres, its generation require infeasible measurement procedure in vivo.

Finally, Tuch tried to reduce the acquisition time by arranging the samples on spherical shell in **q**-space. The signals sampled on two-dimensional spherical coordinates preserve the high angular resolution property and allows model-free and more complicated methods to recover ODF or fODF. Next chapter, three types of methods working with HARDI will be discussed in detail.

# Chapter 3

# Methods for HARDI

HARDI samples diffusion signals on the surface of a sphere in **q**-space whose radius is determined by selected **q** value. It preserves the high angular resolution property of DSI but does not have the heavy measurement burden as DSI. Meanwhile it does not require any format of model assumption like DTI, and is adaptive to more complicated processing methods. In this chapters, we generally categorize HARDI related methods into three groups: multicompartment model, QBI, and SD. They differ from each other in modelling, methodologies, and results.

## 3.1 Multicompartment Model

### 3.1.1 Discrete Mixture of Gaussian Diffusion Model

Diffusion tensor imaging can be used to analyse the apparent water diffusion with the assumption of Gaussian diffusion model. After fitting the diffusion weighted signals to diffusion tensor, multiple useful measures can be extracted from the tensor. The major eigenvector infers the orientation of underlying fibre, and other measures as by-products can also indicate the properties of diffusivities. But limited by the simplicity of it model assumption, DTI can only yield the correct estimation of single fibre existing in a voxel of white matter region, but not multiple fibres.

An instinct idea to overcome this main limitation of DTI is to complicate the single tensor model with a discrete mixture of multiple Gaussian tensor models [10, 11]. Under

the Gaussian diffusion model, the normalized diffusion signal in a voxel can be written as [4, 10, 5, 7],

$$E(\mathbf{q}) = e^{-\mathbf{q}^T \mathbf{D} \mathbf{q} \Delta}, \tag{3.1}$$

where $\mathbf{q}$ is the diffusion gradient wave-vector $\mathbf{q} = \gamma \delta \mathbf{G}$, $\gamma$ is the gyromagnetic ratio, $\delta$ is the diffusion gradient duration, $\mathbf{G}$ is the diffusion gradient, $\Delta$ is the effective diffusion time, and $\mathbf{D}$ is the diffusion tensor. Before giving the multiple compartment model, a several assumptions need to be stated [10]. First of all, the inhomogeneity consists of a discrete number of homogeneous regions. Second assumption statement is the distinct regions are separated by a greater distance than the diffusion mixing length which can guarantee a slow exchange between regions. The last assumption is the diffusion signal of every neural fibre fits Gaussian model. Supported by aforementioned assumptions, a finite mixture of Gaussian diffusion tensor can be given as,

$$E(\mathbf{q}) = \sum_i f_i e^{-\mathbf{q}^T \mathbf{D_i} \mathbf{q} \Delta}, \tag{3.2}$$

where $f_i$ is the volume fraction of each distinct fibre. To explain the acquired diffusion signal, the unknown parameters ($f_i$ and $D_i$) have to be estimated. Owing to the complexity of unknown parameters, conventional expectation maximization can not be used to solve this Gaussian mixture problem. The authors in [10] deployed the technique of gradient descent which is capable to resolve the underlying fibres. Another work in [63] deployed Markov chain Monte Carlo technique to estimate the parameters of single fibre model as well as two fibres models separately and compared their probabilities.

The discrete mixture of Gaussian diffusion model performs better than DTI in the white matter regions which accommodate multiple heterogeneous fibres. But the usage of mixture model requires non-linear fitting which is often unstable. Thus, multiple restarts of optimization method are needed to preventing the algorithm converging to local minimum. Furthermore the accuracy of unknown parameter estimation highly relies on the determination of the number of tensors. Besides those, analysis in [49] argues that a large number diffusion images are required which notably increases acquisition time.

### 3.1.2 Diffusion Basis Functions Decomposition

Although the discrete mixture of Gaussian models can resolve multiple fibres in a voxel, it requires non-linear fitting procedure which is painful. Right after its proposal, an effort had been made to convert the non-fitting into linear fitting by deploying diffusion basis functions to represent diffusion signals [49]. The author of this paper started the analysis

with discrete mixture of Gaussian models. But what is interesting is that, his final signal model is coincidentally the same as the discretized version of SD (discussed in Section 3.3). Moreover this paper introduced sparsity constraint on results which is also used in SD methods [40, 44, 65] as well as in our algorithm in this thesis.

This method depends on the assumption of Gaussian mixture model (GMM) same as (3.2). As what has been done in [10], the eigenvalues are fixed by either estimation or priori information in order to reduce the complexity of computation. Based on the particularly chosen eigenvalues, a fixed set of diffusion basis functions (DBF), $\{\Phi_i\}$, can be determined to represent diffusion signals. The orientations of DBF are distributed uniformly on unit sphere. Accordingly, the normalized diffusion signal can be expressed now as,

$$E(\mathbf{q}) = \sum_{i=1}^{N} \alpha_i \Phi_i(\mathbf{q}) + \varepsilon, \tag{3.3}$$

where $\mathbf{q}$ is the diffusion gradient wave-vector, $\alpha_i \geq 0$ is the coefficient of corresponding DBF, and $\varepsilon$ is the typical Rician distributed noise. To be specific, the $i$th DBF is defined as,

$$\Phi_i(\mathbf{q}) = e^{-\mathbf{q}^T \mathbf{D}_i \mathbf{q} \Delta}, \tag{3.4}$$

with $\Delta$ as the diffusion time and $\mathbf{D}_i$ as the $i$th diffusion tensor. Now the objective is to determine the values of DBF coefficients ($\alpha_i$) to reproduce the measured signal. Under the linear frame, (3.3) can be discretized into a matrix multiplication format,

$$E = \mathbf{\Phi}\alpha + \varepsilon. \tag{3.5}$$

To make the estimation of $\alpha_i$ stable and robust, prior information is added so as to pursue the best values. Basis Pursuit technique was used to creating a linear programming problem as below,

$$\begin{aligned} \min \quad & ||\alpha||_1 = \sum_i \alpha_i = \mathbf{1}^T \alpha \\ \text{subject to} \quad & \mathbf{\Phi}\alpha = E,\ \alpha_i \geq 0, \forall i \end{aligned} \tag{3.6}$$

where $\mathbf{1}$ is a vector with all its components to be one. The regularization terms including $\ell_1$-norm $||\alpha||_1$ reinforcing the sparsity of coefficients and non-negativity constraint guaranteeing reasonable results. In the same year later, another independent researcher Yap in [40] gave the same format of optimization problem.

Besides sparsity and non-negativity regularizers, the authors in [49] also suggested another two regularizers to emphasize the sparsity and spatial piece-wise smoothness. For an arbitrary voxel location $r \in \mathbb{R}^3$, $\alpha_{ir}$ denotes the $\alpha_i$th coefficient at the $r$ voxel location

and $\mathcal{N}_r$ stands for the second order spatial neighbourhood of $r$: $\mathcal{N}_r = s : |r - s| < 2$. Thus the difference between the coefficients of $r$ and those of neighbours' can be expressed as,

$$U_s(\alpha, r) = \sum_{s:s\in\mathcal{N}_r} \sum_i w_{irs}(\alpha_{ir} - \alpha_{is})^2. \tag{3.7}$$

with the control of anisotropic weight factors $w_{irs} = (s - r)^T \mathbf{D}_i(s - r)/||s - r||^4$. We can minimize $U_s(\alpha, r)$ to penalize the difference between coefficients of neighbours. In addition, sparsity constraint can be added not in the format of $\ell_1$-norm but by forcing each $\alpha_{ir}$ coefficient to be different from their arithmetic mean, $\bar{\alpha}_r = \sum_i \alpha_{ir}/N$. To be specific, minimizing $U_c(\alpha, r) = -\sum_i(\alpha_{jr} - \bar{\alpha}_r)^2$ can force $\alpha$ to be sparse. Including these two regularizers, the second cost function of optimization problem proposed in [49] is

$$U(\alpha, r) = ||E - \Phi\alpha_r||_2^2 + \mu_s U_s(\alpha, r) + \mu_c U_c(\alpha, r). \tag{3.8}$$

Although the analysis of diffusion signal started with GMM in this paper, the final formats of optimization problems actually is the discretized version of spherical deconvolution of HARDI which we will discuss in Section 3.3. The main contribution of [49] is to introduce three important regularization terms: sparsity, non-negativity and spatial smoothness. In the following related spherical deconvolution research, a number of papers implemented sparsity constraint [40, 65, 66, 67]. On the other side, according to [50], the spatial smoothness regularizer suggested here is so strong as to reduce the angular resolution of reconstructed fODF.

## 3.2 QBI and Related Methods

Unlike multicompartment model deriving from the idea of Gaussian diffusion model (DTI), the QBI methods do not depend on any format of diffusion model assumption. Actually the idea came from the generation of ODF which was firstly defined in Tuch's DSI papers [60, 11]. In these papers, EAP was recovered from DSI method and then projected on to unit sphere to generate its marginal distribution ODF. DSI suffers from its heavy measurement requirement which is completely unnecessary to generate EAP's marginalization ODF. Two-dimensional results (ODF) only require two-dimensional measurements (HARDI).

### 3.2.1 QBI

Q-ball imaging (QBI) was firstly proposed by Tuch in [24] as an alternative solution to resolve intra-voxel fibre crossing with HARDI signals. As opposed to mixture model decomposition of HARDI, QBI is a model-independent method which does not reply on

the assumption of Gaussianity or multi-Gaussianity. Instead, the method deploys Funk-Radon transform (FRT), aka the spherical Radon transform, to estimate ODF which is the marginal distribution function of EAP. ODF is not a complete description of diffusion propagator but yields the orientational structure of the diffusion propagator.

If simplify (2.16) and denote EAP with $P(\mathbf{r})$, then the relationship between $P(\mathbf{r})$ and normalized diffusion signal, $E(q)$, can be given by

$$P(\mathbf{r}) = \mathcal{F}(E(\mathbf{q})) \tag{3.9}$$

where $\mathbf{r}$ is the spin displacement vector and $\mathcal{F}$ represents the three-dimensional Fourier transform with respect to diffusion wavevector $\mathbf{q}$ [24, 30, 64]. The diffusion ODF $\psi(\mathbf{u})$ is defined as the radial projection of the diffusion function [11, 30, 24, 26],

$$\psi(\mathbf{u}) = \frac{1}{Z}\int_0^\infty P(r\mathbf{u})dr, \tag{3.10}$$

where $\mathbf{u} \in \mathbb{S}^2$ and Z is a normalization constant. The definition of ODF in (3.10) reveals its mathematical nature, a two-dimensional marginal distribution function defined on sphere. ODF quantifies the probability with which water molecules undergo displacement along various spatial directions [23, 24] during the mixing time. Furthermore, the definition also indicates that ODF can be derived from diffusion propagator by radial projection which is not an efficient way owing to a number of limitations [24]. To simplify the calculation of ODF, the author in [24] suggested to reconstruct ODF directly from HARDI signals sampled on a single shell in $\mathbf{q}$-space.

FRT is proposed in [68], which is defined as a transform from the sphere to sphere. For an arbitrary spherical function $f(\mathbf{u})$, $\mathbf{u} \in \mathbb{S}^2$, the FRT $\mathcal{G}$ is defined as the integral of $f(\mathbf{u})$ over the equator which is perpendicular to $\mathbf{u}$, as below,

$$\mathcal{G}[f(\mathbf{u})](\mathbf{w}) = \int_{\mathbf{u}\in\mathbf{w}^\perp} f(\mathbf{u})d\mathbf{u} = \int f(\mathbf{u})\delta(\mathbf{u}^T\mathbf{w})d\mathbf{u}, \tag{3.11}$$

where $\delta$ is Dirac delta function. Extending FRT to an three-dimensional function $f(\mathbf{x})$, a new definition can be written as,

$$\mathcal{G}[f(\mathbf{x})](\mathbf{w}, r') = \int f(\mathbf{x})\delta(\mathbf{u}^T\mathbf{w})\delta(|\mathbf{x}| - r')d\mathbf{x}. \tag{3.12}$$

Simplifying the notation of FRT, the transform in (3.12) here is denoted as $\mathcal{G}_{r'}$. Derived from the previous research of Tuch's in [23], a strong approximation of ODF can be obtained from FRT of diffusion signal, which is,

$$\psi(\mathbf{u}) \approx \frac{1}{Z}\mathcal{G}_{q'}(E(\mathbf{q})), \tag{3.13}$$

where $Z$ is normalization constant and $q'$ is the radius of single sphere shell where signals are sampled. (3.13) only gives an approximation of ODF. A detailed expression of ODF based on FRT is given in [24],

$$\psi(\mathbf{u}) = \mathcal{G}_{q'}(E(\mathbf{q})) = 2\pi q' \int P(r,\theta,z) J_0(2\pi q' r) dr d\theta dz, \tag{3.14}$$

where $J_0$ is the zeroth-order Bessel function and $P(r,\theta,z)$ is the cylindrical coordinates of $P(\mathbf{r})$. (3.14) reveals that the Funk-Randon transform of diffusion signals can yield ODF which is the integral of EAP over range variable. According to (3.14), the calculation of ODF requires equator integral of diffusion signals. But the signal sampling scheme does not always coincide with the points on equators. Therefore some spherical interpolation algorithms can be employed to the grid sampling points. Furthermore, the computation of QBI can be simplified using matrix multiplication as the form $\boldsymbol{\psi}= \mathbf{Ae}$, where $\mathbf{A}$ is FRT reconstruction matrix and $\mathbf{e}$ stands for signal vector. Overall, QBI is a successful method which can resolve multiple intravoxel fibres with HARDI. Its advantages include model independence, linearity, relatively easy implementation, and medium measurement requirement.

### 3.2.2 QBI with Constant Solid Angle

QBI is a successful high angular resolution imaging technique which can resolve multiple intravoxel fibre orientations. Following the proposal of standard QBI, a number of methods were proposed aiming to turn it to be fast, robust and reliable.

Authors in [29, 64] focus on the definition of ODF in (3.14) which, they believed, is not the real marginal distribution of EAP. Apparently, the ODF defined in (3.14) is just a linear radial projection of EAP, which does not account for the quadratic growth of the volume element with respect to its distance from the origin. The consequence of this neglect is the product of nondistribution functions, and the need for artificial postprocessing. To overcome this deficiency, a new definition of ODF is suggested in [29] with constant solid angle (CSA). Being consistent with the notations in Q-ball imagining, $P(\mathbf{r})dv$ stands for the displacement probability of a molecule from origin to the infinitesimal volume $dv$. If we use $(r,\theta,\phi)$ to parametrize cylindrical coordinates, the volume element can be expanded as $dv = r^2 dr d\Omega$, while $d\Omega = \sin\theta d\theta d\phi$ is the infinitesimal solid angle element. Therefore the probability of diffusion in direction $\mathbf{u}$ through solid angle $d\Omega$ can be computed by the marginalization of $P(\mathbf{r})$ as below,

$$\psi(\mathbf{u}) d\Omega = \int_0^\infty P(r\mathbf{u}) r^2 dr d\Omega. \tag{3.15}$$

Eliminating solid angle $d\Omega$ from both sides of the above equation, a new definition ODF can be simplified as [29, 64]

$$\psi(\mathbf{u}) = \int_0^\infty P(r\mathbf{u})r^2 dr. \tag{3.16}$$

Compared with (3.10), the only difference in (3.16) is the factor $r^2$ which allows the new definition to yield normalized and dimensionless ODF. Without the factor $r^2$, the computation of ODF puts a artificial weight on $P(\mathbf{r})$ which varies according to the value of $1/r^2$.

Furthermore, assuming the $P(\mathbf{r})$ follows a standard Gaussian distribution as [29],

$$P(\mathbf{r}) = \frac{1}{(2\pi)^{\frac{3}{2}}|\mathbf{D}|^{\frac{1}{2}}} e^{-\frac{1}{2}\mathbf{r}^{-T}\mathbf{D}^{-1}\mathbf{r}} \tag{3.17}$$

where $\mathbf{D}$ is the diffusion tensor. The ODF estimated by the new definition (3.16) can be expressed as [29],

$$\psi(\mathbf{u}) = \frac{1}{4\pi|\mathbf{D}|^{\frac{1}{2}}((\mathbf{u}^T\mathbf{D}^{-1}\mathbf{u})^{\frac{3}{2}}}. \tag{3.18}$$

### 3.2.3 QBI with Spherical Harmonic Basis

The reconstruction of original QBI in [24] requires spherical interpolation and employs spherical radial basis functions. As a matter of fact, a number of spherical basis sets can be implemented to accomplish this mission. Spherical harmonics (SH) functions are chosen in [26, 27] which can bring a certain number of advantages, such as analytic solution of the reconstruction and interpretation of results in frequency domain.

The SH based representation of diffusion signals $E(\mathbf{q})$ can be written as a linear combination of spherical harmonic basis functions $Y_l^m(\mathbf{u})$,

$$E(\mathbf{q}) = \sum_{l=0}^{L}\sum_{m=-l}^{l} s_l^m Y_l^m(\mathbf{u}), \tag{3.19}$$

with $s_l^m$ denoting the spherical harmonic series coefficients and L being the harmonic series order. Using $(\theta, \phi)$ as the coordinates on sphere, the basis function is analytically expressed as [26, 27],

$$Y_l^m(\mathbf{u}) = Y_l^m(\theta,\phi) = (-1)^m \left(\frac{2l+1}{4\pi}\right)^{\frac{1}{2}} \left(\frac{(l-m)!}{(l+m)!}\right)^{\frac{1}{2}} P_l^m(\cos\theta)e^{im\phi}, \tag{3.20}$$

where $P_l^m(x)$ is the associated Legendre polynomials:

$$P_l^m = \frac{(1-x^2)^{m/2}}{2^l l!}\frac{d^{l+m}}{d^{l+m}x}(x^2-1)^l. \tag{3.21}$$

Similar with Fourier transform of signals, the expansion of diffusion signals in (3.20) is the frequency representation of signals on sphere. To achieve this format, the coefficient $s_l^m$ need to be determined first. Since the basis function $Y_l^m(\mathbf{u})$ are orthogonal to each other,

$$\int_{\theta=0}^{\pi}\int_{\phi=0}^{2\pi} Y_l^m(\mathbf{u})Y_{l'}^{m'*}(\mathbf{u})d\phi d\theta = \frac{4\pi}{2l+1}\delta(l-l')\delta(m-m'), \tag{3.22}$$

where $\delta(x)$ is Dirac delta function, we can integrate $E(\mathbf{q})Y_l^{m*}(\mathbf{u})$ over the sphere to determine the value of each coefficient $s_l^m$. Theoretically, the complete representation requires basis functions of all orders, but in practise $L$ is limited by a constant number which turn (3.19) into an approximation of diffusion signals. The accuracy of this approximation is controlled by the choice of highest order $L$. Detailed experiments are conducted in [26], which illustrates the trade-off between high and low $L$ values.

After expressing diffusion signals in the format of SH basis functions, according to (3.12) and (3.19), the FRT of $E(\mathbf{q})$ can be rewritten as [26]

$$\psi(\mathbf{u}) = \sum_{l=0}^{L}\sum_{m=-l}^{l} c_l^m \oint_{\mathbf{q}\perp\mathbf{u}} Y_l^m(\mathbf{q})d\mathbf{q}. \tag{3.23}$$

The above equation changes the integral of diffusion signals into the integral of SH basis functions. The desirable property of SH is that its sphere equator integral has a simple format, specifically [26, 27],

$$\frac{1}{2\pi}\oint_{\mathbf{q}\perp\mathbf{u}} Y_l^m(\mathbf{q})d\mathbf{q} = P_l(0)Y_l^m(\mathbf{u}), \tag{3.24}$$

where $P_l(x)$ is Legendre polynomial of order $l$ and $\mathbf{u}\in\mathbb{S}^2$. Substituting this analytic integral solution of $Y_l^m(\mathbf{u})$ into (3.23), we can obtain the analytic representation of estimated ODF based on SH basis functions [26, 27],

$$\psi(\mathbf{u}) = \sum_{l=0}^{L}\sum_{m=-l}^{l} 2\pi c_l^m P_l(0)Y_l^m(\mathbf{u}). \tag{3.25}$$

Implied by the aforementioned equations, the estimation of ODF can be obtained by three steps [26]: (1) SH decomposition of diffusion signals, (2) integrate SH basis function by multiplying Legendre polynomials $P_l(0)$, and (3) synthesize modified SH basis functions along with corresponding coefficients $c_l^m$.

### 3.2.4 Limitation of QBI

The nature of single-shell acquisition inherent in HARDI imposes constraints on the angular resolution of estimated ODFs, with higher values of the diffusion scintillation parameter $b$-value leading to better resolvability between various diffusion modes within a given voxel. This fact is exemplified in Fig. 3.1, which shows a simulated ODF[1] (top subplot) corresponding to two fibre tracts crossing each other at an angle of 60 °. At the same time, the bottom row of subplots depict the ODFs which have been recovered from the associated HARDI data generated with $b \in \{1000, 3000, 5000\}$ s/mm$^2$. One can see that the best angular resolution is attained at the maximum value of $b$= 5000 s/mm$^2$, as expected. It is also worthwhile noting that the above effect is intrinsic in both the FRT [24] and solid angle [29, 64] formulations of QBI.

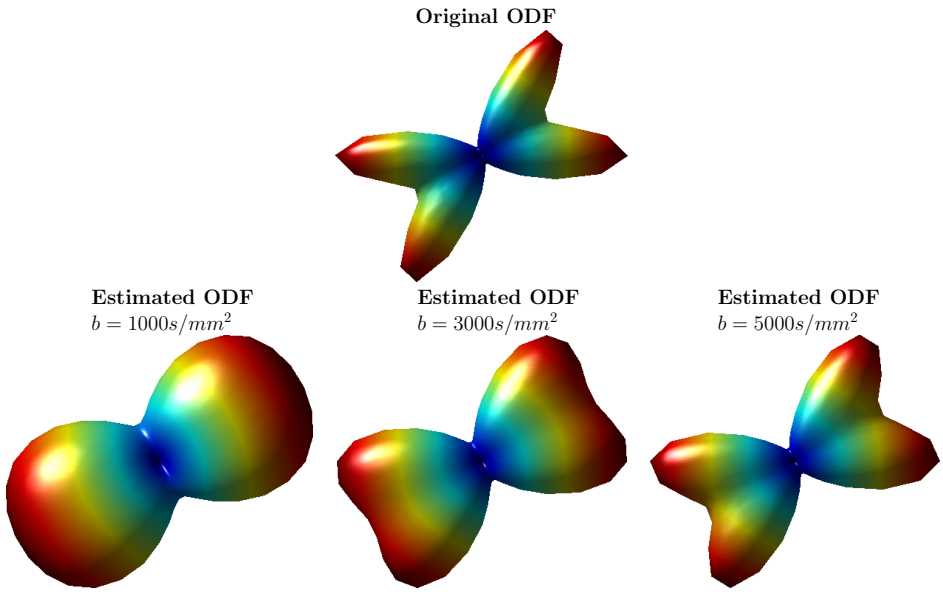


Figure 3.1: (Top) Original ODF; (Bottom) Estimated ODFs obtained from HARDI data generated with (left to right) $b$=1000, 3000, and 5000 s/mm$^2$.

As tempting as it might seem at the first glance, working with relatively high $b$-values is usually avoided in practice for two main reasons. First, an increase in the $b$-value is typically achieved through using longer mixing times, which unavoidably leads to substantially low

[1]The ODF was generated using a standard Gaussian mixture model [25] with equal volume fractions, FA=0.8 and MD = $7 \cdot 10^{-4}$ mm$^2$/s.

values of signal-to-noise ratio (SNR) [69]. Second, using higher $b$-values makes the diffusion data less sensitive to the effects of fast diffusion, which is often associated with the concept of "free water" [70], and hence represents a diagnostically important diffusion regime. As a result, diffusion data are commonly acquired with relatively low values of $b$, typically around 1000 s/mm$^2$. Needless to add, in such cases, poor directional resolution of ODFs may become an issue, especially for such applications as fibre tractography [31].

## 3.3 Spherical Deconvolution of HARDI

Although in some SD methods Gaussian diffusion model is adopted as single fibre response (SFR), the usage of DTI model is not mandatory in SD. Meanwhile, SD of HARDI does not use the Fourier relationship between EAP and normalized diffusion signals as DSI and QBI. SD methods reply on the model assumption that normalized HARDI signals can be modelled as the convolution of two spherical functions: single fibre response and fibre orientation distribution function. Due to its inverse problem nature, SD method normally requires additional priori to constrain the results. In this section, SD methods are presented with a variety of typical regularization constraints, such as sparsity, non-negativity, Tikhonov regularization, and spatial regularizations.

### 3.3.1 Filtered Spherical Deconvolution

DTI is successful as it can recover the orientation of single fibre with the major eigenvector of diffusion tensor and generate other contrasts, such as FA and RA which contain the information of anisotropy. But its ability is limited with the voxels which contain only one fibre. To overcome DTI's limitation, alternative methods were later proposed to resolve multiple fibres within one voxel, like multicompartment model, $q$-space imaging, DSI, and QBI. But they all have their own limitations. Tournier is the first one who suggested to deploy spherical deconvolution (SD) on HARDI in [33]. The motivation of this new method proposal is to to recover fibre orientation distribution function (fODF) directly from HARDI without prior assumption or estimation about the number of the fibres.

The concept of SD methods replies on several assumptions [33]. Firstly, it is a common assumption that there is no exchange between spatially distinct fibre bundles. Therefore signals emanating from different regions can be assumed to be added independently to generate the total measured signal. Secondly, for curved fibres, there is effectively no exchange between orientationally distinct sections of the same fibre bundle, in which case the

signal can be approximated by the sum of signals emanating from different orientationally distinct regions present in the sample. The last assumption is that all fibres all over the brain share the same diffusion characteristics. In another word, the signal measured from single fibre can be represented as a cylindrically symmetric spherical function $R(\theta)$ with $\theta$ and $\phi$ being spherical coordinates. Based on these assumptions, the HARDI signal $E(\theta, \phi)$ acquired from single voxel accommodating multiple distinct fibres can be approximated as the sum of the response functions of each fibre, weighted by their own volume fractions. Each single fibre has its distinct orientation. The signal $E(\theta, \phi)$ then can be written as,

$$E(\theta, \phi) = \sum_i f_i \mathbf{A}_i R(\theta) \tag{3.26}$$

where $f_i$ is the volume fraction of the $i$th fibre, and $\mathbf{A_i}$ is the rotation operation matrix which rotates the $i$th fibre to direction $(\theta_i, \phi_i)$. (3.26) looks similar with the multicompartment tensor model in (3.2) whose general case is the convolution of single fibre response and fODF $F(\theta, \phi)$ as below,

$$E(\theta, \phi) = F(\theta, \phi) \otimes R(\theta), \tag{3.27}$$

where $\otimes$ stands for spherical convolution. Comparing (3.27) with (3.26), we can tell that fODF $F(\theta, \phi)$ contains the information of each distinct fibre, including fibre orientation and volume fraction. A common case is that we have $N$ fibres in one voxel, where $N$ is a limited natural number, the corresponding fODF is nothing but the sum of $N$ Dirac delta functions on sphere weighted by corresponding volume fractions, which indicate the orientations of related fibres. The most inspiring conclusion we can make from (3.27) is that, once we have acquired $E(\theta, \phi)$ and $R(\theta)$, the unknown part fODF $F(\theta, \phi)$ can be computed by performing spherical deconvolution of $R(\theta)$ from $E(\theta, \phi)$.

To make the convolution computation in (3.27) simple and make it model-independent, a set of spherical and rotational harmonic functions are deployed. Spherical harmonic functions have been introduced in Section 3.2.3 while the rotational harmonic function is another complete orthonormal basis over the space of pure rotations [33]. Each rotational harmonic function is determined by three parameters, a harmonic order $n$ ($n > 0$) and two phase factors $m$ and $l$ ($-n \leq m, l \leq n$). The $n$th order spherical harmonic decomposition $E^n$ of signal $E(\theta, \phi)$ can be estimated using linear least square fit. On the other side of equation, represent $R(\theta)$ with its rotational harmonic decomposition $\mathbf{R}^n$ which is a $(2n + 1)(2n + 1)$ matrix and define $\mathbf{F}^n$, a vector of length $(2n + 1)$, as the $n$th order spherical harmonic decomposition of $F(\theta, \phi)$. The spherical convolution in (3.27) can be rewritten in a matrix format,

$$E^n = \mathbf{R}^n \mathbf{F}^n, \tag{3.28}$$

which indicates that $\mathbf{F}^n$ can be simply recovered by inverting $\mathbf{R}^n$, the operation of spherical deconvolution. Furthermore the computational complexity can be further reduced by considering the symmetry properties of diffusion signal. Finally, there was an effort made to reduce the effect of noise by attenuating or eliminating the high frequency components of $\mathbf{F}^n$. The reduction of noise was build on the cost of angular resolution. Therefore, in the following work [34], the author named this version of SD method as filtered spherical deconvolution (filtered SD) for the low-pass filter effect.

The introduction of SD on HARDI in [33] created a new methodology to analyse HARDI signals. It possesses a number of advantages, such as simple computation, no priori information needed, and model independence. The nature of SD is a inverse problem which is ill-posed. A few problems of SD need to be tackled like being sensitive to noise and the presence of artefactual negative side lobes.

### 3.3.2 Maximum Entropy Spherical Deconvolution (MESD)

A comparative study in [71] shows that the SD method [33] on HARDI is unstable and performs worse than QBI [24]. Therefore the authors in [38] intended to improve the stability of SD method by maximizing the entropy of fODF in (3.27). Maximum entropy methods have been proved to be useful in a number of reconstruction and inverse problems [72], in particular effective with deconvolution.

The new method, maximum entropy spherical deconvolution (MESD), proposed in [38] still replies on the convolution assumption in (3.27). In addition, the authors believe that the fODF, $F(\mathbf{u}), \mathbf{u} \in \mathbb{S}^2$, is supposed to have sharp peaks which align with the directions of underlying fibres. Moreover they suggested to derive a non-linear representation of fODF along with a maximum-entropy augment. Since $F(\mathbf{u})$ is a probability density function, it integrates to one over the sphere domain,

$$\int_{\mathbb{S}^2} F(\mathbf{u}) d\mathbf{u} = 1, \tag{3.29}$$

where $u \in \mathbb{S}^2$ and $\mathbb{S}^2$ denotes the unit sphere in $\mathbb{R}^3$. The information content of the fODF is expressed as,

$$I[F] = \int_{\mathbb{S}^2} F(\mathbf{u}) log F(\mathbf{u}) d\mathbf{u}. \tag{3.30}$$

We can rewrite the convolution model in (3.27) with wave vector $\mathbf{q}$ from $\mathbf{q}$-space and unit vector $\mathbf{u} \in \mathbb{S}^2$ as below,

$$E(\mathbf{q}) = \int_{\mathbb{S}^2} R(\mathbf{q}, \mathbf{u}) F(\mathbf{u}) d\mathbf{u}, \tag{3.31}$$

where $R(\mathbf{q}, \mathbf{u})$ is the single fibre response. The new expression implies that each signal measurement provides a constraint on $F(\mathbf{u})$. Augmenting each constraint (3.31) onto (3.30) with the Lagrange multipliers, we can get cost function as [38],

$$I[F] = \int_{\mathbb{S}^2} \left( F(\mathbf{u}) log F(\mathbf{u}) - F(\mathbf{u}) \sum_{i=1}^{N} (\lambda_i R(\mathbf{q}_i, \mathbf{u})) - F(\mathbf{u})\mu \right) d\mathbf{u}, \tag{3.32}$$

where $\mathbf{q}_i, i = 1, \ldots, N$, are the wave vectors of the signal, the $\lambda_i$ are Lagrange multipliers for the constraints from the data and the Lagrange multiplier $\mu$ controls the normalization of $F(\mathbf{u})$. The minimum value of information content ($I[F]$) can be obtained by taking the variational derivative $\delta I[F]$ and solving $\delta I[F] = 0$. Following the above steps, $I[F]$ is minimum when [38]

$$F(\mathbf{u}) = exp \left( \lambda_0 + \sum_{i=1}^{N} \lambda_i R(\mathbf{q}_i, \mathbf{u}) \right), \tag{3.33}$$

where $\lambda_0 = \mu - 1$. The values of $\lambda_i$ can be further determined by solving [38]

$$\int_{\mathbb{S}^2} F(\mathbf{u}) R(\mathbf{q}_i, \mathbf{u}) d\mathbf{u} = E(\mathbf{q}_i). \tag{3.34}$$

This MESD method improves the performance of original SD method [33] by regularizing the information content of reconstructed fODF. Another contribution of MESD is the non-negativity constraint of fODF which is implicitly included in the computation of entropy. The significant disadvantage of MESD is that its computation time is much greater than original SD method. Moreover, the optimization problem does not guarantee to converge to global minimum.

### 3.3.3 Constrained Spherical Deconvolution (CSD)

When Tournier first time suggested to implement spherical deconvolution on HARDI in [33], one of drawbacks mentioned in that paper was the artefactual negative side lobes. In the following work [34], he explicitly added non-negativity constraint to convolution model to create a new method, constrained spherical deconvolution (CSD), unlike MESD [38] which includes the constraint implicitly.

In the previous work [33], Tournier intended to suppress large spurious negative lobes of reconstructed fODF by attenuating or eliminating the high frequency components. The direct cost of this type of low-pass filter effect is the angular resolution. An alternative way to reduce the effect of noise is to add a constraint on the presence of negative values in the reconstructed fODF. The format of the constraint in [34] was chosen as a modified Tikhonov regularization method.

The CSD method inherits the basis of SD from filtered SD illustrated in (3.28). It takes the results of filtered SD as the initial estimate of fODF, $\mathbf{F}_0^n$ generated with truncated harmonic order $n_{max} = 4$, and then iteratively improve it. In the following steps, for the $i$th iteration, the amplitude $\mathbf{f}$ of fODF is computed along a large set of $N$ directions (uniformly distributed on unit sphere),

$$\mathbf{f} = \mathbf{P}\mathbf{F}_\mathbf{i}^\mathbf{n}, \tag{3.35}$$

where $\mathbf{P}$ maps $\mathbf{F}_\mathbf{i}^\mathbf{n}$ onto the amplitudes $\mathbf{f}$ along $N$ directions. A constraint matrix $\mathbf{L}$ can be determined as:

$$\mathbf{L}_{m,n} = \begin{cases} \mathbf{P}_{m,n} & \mathbf{f}_m < \tau \\ 0 & \mathbf{f}_m \geq \tau \end{cases} \tag{3.36}$$

where $\tau$ is a constant which plays as a user-specified threshold. If the amplitude of $\mathbf{f_m}$ is below $\tau$, it should be suppressed to be zero. Once $\mathbf{L}$ is decided, an iterative process can be started to improve fODF by solving the following optimization problem:

$$\mathbf{F}_i^n = \arg\min\{||\mathbf{R}^n\mathbf{F}^n - S^n||^2 + \lambda^2||\mathbf{L}\mathbf{F}_i^n||^2\} \tag{3.37}$$

where $\lambda$ control the relative weighting between two terms. The first term sticks with the SD model and second term will force the selected components of $\mathbf{F}_i^n$ to be zeros. The iterations will continue until there is no observation of further change in the matrix $\mathbf{L}$. According to the experiments in [34], a typical number of iterations is between 5 and 10.

Within the framework of SD, CSD provides another way to control the effect of noise by penalizing the negative side lobes. It improves the reliability of reconstructed fODF within voxel. This improvement may lead to the development of more robust tractography algorithms. Besides CSD, there are numbers of methods [47, 50, 73, 44] implement non-negativity constraint in their optimization problems.

### 3.3.4 Richardson-Lucy (RL) Spherical Deconvolution

Dell'Acqua started the discussion of HARDI signal representation with multicompartment model in [41] and then found the link between multicompartment model with spherical

deconvolution, indicating that the fODF in (3.27) is nothing but several discrete Dirac delta functions on unit sphere. Further, an significant contribution in this paper is to take account of isotropic diffusion besides the anisotropic diffusion of fODF. To combine both of two components, Dell′Acqua suggested to use a new term, fibre orientation function (FOF), to represent the weights of the anisotropic and isotropic diffusion.

An modified version of SD in (3.27) is given as below to explicitly consider the effect of isotropic diffusion:

$$E(\theta, \phi) = F(\theta, \phi) \otimes R(\theta) + W_{\text{iso}}. \tag{3.38}$$

where $(\theta, \phi)$ and $R(\theta)$ are defined same as in (3.27) and $W_{iso}$ stands for the isotropic component. Similar as the anisotropic component, $W_{iso}$ can be further extended as convolution of single fibre response and an isotropic constant ($w_{iso}$), in which case a generalized FOF ca be defined as:

$$FOF(\theta, \phi) = F(\theta, \phi) + w_{\text{iso}}. \tag{3.39}$$

Therefore, the signal can be regarded as the convolution of FOF and single fibre response,

$$E(\theta, \phi) = FOF(\theta, \phi) \otimes R(\theta), \tag{3.40}$$

from which we can tell that FOF can be recovered using spherical deconvolution. Richardson-Lucy (RL) algorithm, aka expectation maximization algorithm, was used to implement the operation which is an iterative algorithm modified for Gaussian noise. The motivation of this choice comes from it three advantages. Firstly, it is robust to noise. Secondly, it can tolerate the imprecision of system impulse response. Finally it implicitly applies the non-negativity constraint on reconstructed FOF. All those advantages perfectly fit the properties of HARDI signals and single fibre response.

If $\mathbf{F}_i$, a column vector of length $n$, stands for the the values of FOF along $n$ directions in $i$th iteration, and $\mathbf{E}$ is the column-vector of length $m$ represents the signal sampled on sphere, for each iteration, $\mathbf{F}_i$ can be updated according to

$$\mathbf{F}_{i+1} = \mathbf{F}_i \frac{\mathbf{R}^T \mathbf{E}}{\mathbf{R}^T \mathbf{R} \mathbf{F}_i} \tag{3.41}$$

where $\mathbf{R}$, an $m \times n$ matrix, acts as convolution matrix. Since the computation only requires forward convolution, no matrix inversion is needed which reduces the calculation burden. In addition, considering the symmetry properties of signals and problem, vectors and matrices can be halved so as to speed up the algorithm further.

The introduction of RL algorithm to deconvolution problem in this paper renders the new defined FOF robust to noise and non-negative. Unfortunately, thought the author

added isotropic component to fODF so as to create FOF, they did not regularize its behaviour. Furthermore, FOF as a combination of anisotropic and isotropic diffusions, can not really take them apart which make the usage of FOF difficult.

### 3.3.5 Damped Richardson-Lucy (dRL) Spherical Deconvolution

The aforementioned algorithms working on HARDI signals mainly focus on the anisotropic diffusion. A variety of regularizers, such as sparsity, piece-wise smoothness, and non-negativity, have been adopted to reduce the instability effects owing to noise. The neglect of isotropic diffusion limits their usages within pure white matter regions of the brain. In another word, before implementing algorithms, a selection mask need to be generated to indicate the white matter regions. Otherwise, isotropic diffusion existing in grey matter or cerebrospinal fluid (CSF) will mislead the algorithms to yield spurious peaks in fODF or FOF, even the isotropic components only occupy partial volumes. To expand the implementation regions and improve the quality of recovered FOF defined in [41], Dell'Acqua evolved RL spherical deconvolution with an adaptive regularizer to yield damped Richardson-Lucy algorithm (dRL) spherical deconvolution.

Different from (3.41) in RL, the dRL changes the process of updating the $k$th component of FOF in $i$th iteration by [42]

$$[\mathbf{F}_{i+1}]^k = [\mathbf{F}_i]^k \left( \mathbf{1} + [\mathbf{u}_i]^k \left( \frac{[\mathbf{R}^T\mathbf{E} - \mathbf{R}^T\mathbf{R}\mathbf{F}_i]^k}{[\mathbf{R}^T\mathbf{R}\mathbf{F}_i]^k} \right) \right), \tag{3.42}$$

where $\mathbf{u}_i$ is a vector of the length same as $\mathbf{F}_i$, which performs a damping operation on each component of $\mathbf{F}_i$. When $\mathbf{u}_i$ approaches to 1, the algorithm turns back to (3.41). On the other hand, $\mathbf{F}_{i+1} = \mathbf{F}_i$ while $\mathbf{u}_i = 0$. Since the isotropic component only contributes small amplitude to signal, a strong damping can be applied to smaller FOF components. At the same, light or no damping should be applied to major FOF components. To make damping vector adaptive to the values of FOF, damping vector $\mathbf{u}_i$ is defined through a FOF amplitude dependent vector $\mathbf{r}$ [42]:

$$[\mathbf{r}_i]^k = 1 - \frac{[\mathbf{F}_i^\nu]^k}{[\mathbf{F}_i^\nu + \eta^\nu]^k}. \tag{3.43}$$

Based on $\mathbf{r}$, $\mathbf{u}$ can be written as [42],

$$[\mathbf{u}]^k = 1 - \mu[\mathbf{r}]^k, \tag{3.44}$$

where $\mathbf{r}$ controls the level of damping for each component of FOF and $\nu$ is a geometrical parameter describing the profile of the damping curve and how fast the damping turns on and off. In [42], the values of $\mu$ is decided by the standard deviation of HARDI signal $\mathbf{E}$, to be specific,

$$\mu = \max(0, 1 - 4std(\mathbf{E})). \tag{3.45}$$

Simulation and in vivo experiments and results showed that dRL-SD performs beyond the original RL-SD especially with high partial volume of isotropic components. It perfectly preserves the non-negativity constraint and successfully suppresses the spurious orientations caused by noise or isotropic diffusion. Obviously, the determination of parameter values is a drawback of dRL-SD. Moreover, the definition of FOF derived from RL-SD decides that we can not separate the isotropic and anisotropic components of FOF. That will also affect the angular resolution of FOF.

### 3.3.6 CSD with Fibre Continuity (CSD-FC)

Most of the spherical deconvolution methods mentioned above focus on resolving and analysing anisotropic diffusion against the presence of noise and isotropic diffusion in a voxel by voxel pattern. While applying spherical deconvolution, they intentionally reinforce some properties of fibres by a variety of regularizers. Only two papers [49, 47] noticed that there was similarity between fibres from neighbouring voxels, where they deployed a weighted $L_2$ norm regularizer to implement the reinforcement. Fortunately, it was not the only trial which was made to discover the inter-voxel similarity of fibres. Reisert was inspired by the idea of the gradient vector flow [74], and then came up with a new method in [50], constrained spherical deconvolution with fibre continuity (CSD-FC).

The concept of fibre continuity in [50] starts with the analysis of the property of smooth curve. An function, $f(\mathbf{x}, \mathbf{n}) : \mathbb{R}^3 \times \mathbb{S}^2 \to \mathbb{R}$, can be used to describe a number of fibres passing a volume $\mathbf{x} \in \mathbb{R}^3$ with direction $\mathbf{n} \in \mathbb{S}^2$. The neural fibres are typically smooth curves, which can be interpreted as locally straight. The idea can be written as,

$$f(\mathbf{x}, \mathbf{n}) = f(\mathbf{x} + \alpha\mathbf{n}, \mathbf{n}), \tag{3.46}$$

for a small positive $\alpha$. The physical explanation of the above equation is that if the fibre continuity assumption is satisfied, a fibre passing voxel $\mathbf{x}$ with direction $\mathbf{n}$ will also pass voxel $\mathbf{x} + \alpha\mathbf{n}$. Further expanding the right side of above equation with respect to $\alpha$, the fibre continuity condition is formulated as,

$$\mathbf{n} \cdot \bigtriangledown f(\mathbf{x}, \mathbf{n}) = 0, \tag{3.47}$$

where $\bigtriangledown$ is the gradient of $f$ with respect to $\mathbf{x}$ and "$\cdot$" stands for inner product. There is one more step to inject fibre continuity into spherical deconvolution. Reisert defined a spatial regularizer as,

$$R_{FC}(f) = \lambda \iint_{\mathbb{R}^3 \times \mathbb{S}^2} (\mathbf{n} \cdot \bigtriangledown f)^2 d\mathbf{x} d\mathbf{n} \tag{3.48}$$

with $\lambda$, a user determined parameter, controlling the weight. The spatial regularizer, $R_{FC}(f)$, along with non-negativity were used to improved the results of (3.28).

Experiments and results with synthetic and in vivo data in [50] showed that fibre continuity could improve the quality of reconstructed fODF where the assumption was satisfied. Moreover, the results suggested the FC regularizer performed better than the spatial regularizer used in [49, 47]. The limitation of this method also comes from its local straight assumption which can not be always true.

## 3.4 Conclusion

While different in their assumptions, computational requirements, and performances, virtually most of the above-mentioned SD methods share a common drawback. Particularly, all these methods strongly adhere to the assumption on the data signal to be formed as a convolution of a single fibre response and an fODF, thereby explicitly requiring the presence of a fibre tract at the corresponding voxel. However, situations are frequent in which the data may alternatively be associated with either grey matter or cerebrospinal fluid (CSF), as well as with partial volume fraction voxels containing a mixture of different types of cerebral tissues. In such cases, the diffusion signals should be expected to have a non-negligible isotropic component, which comes at odds with the assumption on fODFs to be sparse or of maximal entropy or even of continuity. A new SD method is needed to take isotropic diffusion into account. Meanwhile it should overcome the side effects of the presence of noise and isotropic diffusion.

# Chapter 4

# Main Contribution

This chapter introduces the proposed SD method in this thesis, spatially constrained sparse deconvolution (SCSD), which is named after the regularization constraints used in it. It explicitly considers the presence of isotropic diffusion and spatially regularizes both anisotropic and isotropic components separately. The method is then formulated as a convex optimization problem and solved with ADMM which yields analytic solutions. Portions of this chapter have been presented at a conference, SPIE Medical Imaging 2014, [67]. The SD method introduced in this chapter derives from early proposed SD method in [67], which is named Min-TV-$\mathcal{L}_1$. Some text and equations from that conference paper are adapted and extended in this chapter.

## 4.1 Data Formation Model

To fix the ideas, we start with a formal setting, in which HARDI data are assumed to be collected over a bounded (open) subset of $\Omega \in \mathbb{R}^3$. In this case, for each $\mathbf{r} \in \Omega$, a HARDI signal $s(\mathbf{u}|\mathbf{r})$ can be viewed as a positive-valued, spherical function $s(\cdot|\mathbf{r}) : \mathbb{S}^2 \to [0, \infty)$, with $\mathbb{S}^2$ denoting the unit sphere in $\mathbb{R}^3$, in which case $\mathbf{u} \in \mathbb{S}^2$ is interpreted as the direction of diffusion gradients. When a given voxel of interest supports a single neural fibre, the corresponding EAP can be closely approximated by a unimodal Gaussian density, in which case the HARDI signal $s(\mathbf{u}|\mathbf{r})$ can be described as [4, 6, 67]

$$s(\mathbf{u}|\mathbf{r}) = s_0(\mathbf{r}) \exp\{-b\ \mathbf{u}^T D(\mathbf{r})\mathbf{u}\}, \quad \forall \mathbf{r} \in \Omega, \tag{4.1}$$

where the diffusion tensor $D(\mathbf{r})$ encodes the directivity and ellipticity of the pattern of local diffusion, while the $b$-value is typically set in the range between 1000 and 3000 s/mm$^2$. Note

that the $b_0$-image $s_0(\mathbf{r})$ is usually acquired through additional measurements and used to normalize the HARDI signals. For the sake of notational convenience, in what follows, the signal $s(\mathbf{u}|\mathbf{r})$ will be assumed to be normalized, implying $s_0(\mathbf{r}) = 1$, for all $\mathbf{r} \in \Omega$.

Although standard in DTI, the model of (4.1) is not applicable in situations when a voxel of interest supports multiple neural bundles [10]. If this is the case, then under some fairly general conditions [25], the HARDI signal $s(\mathbf{u}|\mathbf{r})$ can be assumed to be formed as a (linear) superposition of several "DTI signals" weighted by their respective partial volume fraction coefficients. To formalize such a signal formation model, it is common to use the notion of an SFR [33], which is, in fact, equal to the elementary "DTI signal" given by (4.1). Moreover, in the SD literature, for the definition of a SFR $h(\mathbf{u})$ it is standard to use a cylindrically symmetric diffusion tensor $D_0 = \text{diag}\{\lambda_\perp,\ \lambda_\perp,\ \lambda_\parallel\}$ (with $\lambda_\parallel > \lambda_\perp$), which seems to naturally comply with an expected (local) geometry of neural fasciculi. In such a case, it is straightforward to show that the SFR becomes a zonal (spherical) function [75] that is formally given by [67]

$$h_{\mathbf{v}_0}(\mathbf{u}) = h(\mathbf{u} \cdot \mathbf{v}_0) = \alpha \exp\{-\beta(\mathbf{u} \cdot \mathbf{v}_0)\}, \tag{4.2}$$

with the dot standing for the standard (Euclidean) dot product in $\mathbb{R}^3$, $\alpha \triangleq \exp\{-b\lambda_\perp\}$, $\beta \triangleq b(\lambda_\parallel - \lambda_\perp)$, and $\mathbf{v}_0 = [0,\ 0,\ 1]^T$ denoting the north pole of $\mathbb{S}^2$.

It is worthwhile noting that, due to the property of the SFR in (4.2) to be zonal, its value at a given $\mathbf{u}$ depends only on the angle between $\mathbf{u}$ and a fixed direction $\mathbf{v}$ (e.g., $\mathbf{v} = \mathbf{v}_0$, as in (4.2)), which makes $h_\mathbf{v}(\mathbf{u})$ invariant under rotations around $\mathbf{v}$. This invariance allows the result of convolution with $h_\mathbf{v}(\mathbf{u})$ to be expressed as a function of $\mathbb{S}^2$ (rather than of the orthogonal group $SO(3)$ [32]), in which case our signal formation model becomes [67]

$$s(\mathbf{u}|\mathbf{r}) = \int_{\mathbb{S}^2} h_\mathbf{u}(\mathbf{v}) d\mu(\mathbf{v}|\mathbf{r}) = \int_{\mathbb{S}^2} h(\mathbf{u} \cdot \mathbf{v}) d\mu(\mathbf{v}|\mathbf{r}), \quad \forall \mathbf{r} \in \Omega. \tag{4.3}$$

Here $\mu(\mathbf{u}|\mathbf{r})$ is a probability measure that is used to model the fibre probability distribution over $\mathbb{S}^2$ [46]. In particular, at any $\mathbf{r} \in \Omega$, $\mu(\cdot|\mathbf{r}) : \mathcal{B} \to [0, \infty]$ quantifies the relative frequency of specific fibre orientations over a given element of the Borel sigma algebra $\mathcal{B}$ of $\mathbb{S}^2$.

A mathematically elegant and physiologically meaningful way to interpret the structure of $\mu$ was recently described in [46], where the authors took advantage of the Lebesgue's decomposition theorem to represent $\mu$ as a sum of three components (*viz.*, discrete, absolutely continuous, and singular continuous), each of which is able to model a distinct characteristic of the fibre orientation distribution. In this thesis, we proceed under a simplified assumption on $\mu$ to be absolutely continuous, in which case it can be described in terms of a non-negative, Borel measurable function $f(\mathbf{u}|\mathbf{r})$ as $d\mu(\mathbf{v}|\mathbf{r}) = f(\mathbf{v}|\mathbf{r})d\eta(\mathbf{v})$, with

$\eta$ being the Haar measure of $\mathbb{S}^2$. We note that this simplification seems to be reasonable, considering the fact that, in practical computations, both measurements and resulting estimates are always bounded in value and discrete.

The above simplifying assumption leads to the standard (forward) model for $s(\mathbf{u}|\mathbf{r})$ which reads [35, 40, 42, 46, 44, 67]

$$s(\mathbf{u}|\mathbf{r}) = \int_{\mathbb{S}^2} h(\mathbf{u}\cdot\mathbf{v}) f(\mathbf{v}|\mathbf{r}) d\eta(\mathbf{v}), \quad \forall \mathbf{r} \in \Omega, \tag{4.4}$$

in which case the density $f(\mathbf{u}|\mathbf{r})$ is conventionally referred to as a fibre orientation distribution function (fODF). It should also be noted that the models in (4.3) and (4.4) are stationary, since the SFR is assumed to be fixed within a given voxel as well as across the whole image domain $\Omega$. While only approximative [76], the stationary SD model has nevertheless shown to yield useful reconstructions, while offering the important advantage of tractability and amenability to numerical computations.

Additionally, the structure of fODF $f(\mathbf{u}|\mathbf{r})$ in (4.4) deserves a special consideration. Since the measured diffusion signal receives contributions from both coherently ordered axonal fascicles as well as from their complex and more heterogeneous extra-axonal surroundings (containing astrocytes, glia, and randomly oriented extracellular matrix molecules) [16], it seems reasonable to consider $f(\mathbf{u}|\mathbf{r})$ to be composed of two main terms, *viz.* anisotropic and isotropic. Specifically, following the line of ideas advocated in [42, 46, 44], we model $f(\mathbf{u}|\mathbf{r})$ according to [67]

$$f(\mathbf{u}|\mathbf{r}) = (1 - p_{\text{iso}}(\mathbf{r})) f_a(\mathbf{u}|\mathbf{r}) + p_{\text{iso}}(\mathbf{r}) f_{\text{iso}}(\mathbf{r}), \tag{4.5}$$

with $f_a(\mathbf{u}|\mathbf{r})$ and $f_{\text{iso}}(\mathbf{r})$ representing the anisotropic and isotropic components of the fibre probability distribution, respectively, and $0 \leq p_{\text{iso}}(\mathbf{r}) \leq 1$ controlling their partial volume fractions at $\mathbf{r} \in \Omega$. Moreover, $f_{\text{iso}}(\mathbf{r})$ can be further represented as $f_{\text{iso}}(\mathbf{r}) = \exp(-b\lambda_{\text{iso}}(\mathbf{r}))$, where $\lambda_{\text{iso}}(\mathbf{r}) > 0$ is the apparent diffusivity of the isotropic component [46, 44].

Since the isotropic components $f_{\text{iso}}(\mathbf{r})$ in (4.5) is dissociated from the anisotropic component $f_a(\cdot|\mathbf{r})$, it would no longer be correct to regard $f(\mathbf{u}|\mathbf{r})$ as an fODF. To overcome this notational inconsistency, from now on, we will use this term when referring to $f_a(\mathbf{u}|\mathbf{r})$ instead. At the same time, since $f_{\text{iso}}(\mathbf{r})$ can be viewed as a scalar-valued function of $\mathbf{r} \in \Omega$, it will be referred hereinafter to as an isotropic diffusion map (IDM).

Finally, instead of trying to recover $f_a(\mathbf{u}|\mathbf{r})$ and $f_{\text{iso}}(\mathbf{r})$ along with $p_{\text{iso}}(\mathbf{r})$, we estimate the weighted quantities $f'_a(\mathbf{u}|\mathbf{r}) \triangleq (1 - p_{\text{iso}}(\mathbf{r})) f_a(\mathbf{u}|\mathbf{r})$ and $f'_{\text{iso}}(\mathbf{r}) \triangleq p_{\text{iso}}(\mathbf{r}) f_{\text{iso}}(\mathbf{r})$, such that $f(\mathbf{u}|\mathbf{r}) = f'_a(\mathbf{u}|\mathbf{r}) + f'_{\text{iso}}(\mathbf{r})$, $\forall \mathbf{r}$. We note that, when normalized by $\int_{\mathbb{S}^2} f(\mathbf{u}|\mathbf{r}) d\eta(\mathbf{u})$, $f'_a(\mathbf{u}|\mathbf{r})$

and $f'_{\text{iso}}(\mathbf{r})$ acquire the "flavour" of posterior probabilities. Thus, for example, the values of $f'_a(\mathbf{u}|\mathbf{r})$ are not only indicative of the orientations of neural fibres, but also reflect one's level of confidence that the fibres are actually present at a given location $\mathbf{r}$ at the first place. We support this concept through an experimental study and argue that using the weighted densities $f'_a(\mathbf{u}|\mathbf{r})$ and $f'_{\text{iso}}(\mathbf{r})$ can benefit a number of related applications, such as probabilistic fibre tractography [31].

## 4.2 Spatially Constrained Sparse Deconvolution

### 4.2.1 Model Discretization

As usual, the formalism of matrix-vector multiplications turns out to be the most convenient for formulation of practical solutions. To this end, we first note that (normalized) HARDI data typically consists of a set of diffusion-weighted scans acquired for $K$ directions $\{\mathbf{u}_k\}_{k=1}^K$ of diffusion gradients (with $\mathbf{u}_k \in \mathbb{S}^2, \forall k$). For the convenience of exposition, we concatenate these scans into $K$ row vectors of length $I$, with $I$ being equal to the number of spatial samples (voxels) within $\Omega$. Further, the vectors thus obtained can be organized as the rows of a $K \times I$ matrix $s$, in which case the columns of $s$ correspond to diffusion measurements observed at different spatial locations. In what follows, we use both super- and subscripts to distinguish between the columns $\{s^i\}_{i=1}^I$ and rows $\{s_k\}_{k=1}^K$ of the data matrix $s$.

Discretizing the SD model of (4.4) is the next step. To this end, let $\{\mathbf{v}_j\}_{j=1}^J$ (with $J > K$) be a set of spherical points over which the values of the fODFs are to be recovered. Then, given estimates[1] $\tilde{\alpha}$ and $\tilde{\beta}$ of the SFR parameters in (4.2), we define the $(k, j)$-th element of a $K \times J$ matrix $H = \{h_{k,j}\}$ according to [67]

$$h_{k,j} = \tilde{\alpha} \exp\{-\tilde{\beta}(\mathbf{u}_k \cdot \mathbf{v}_j)\}. \tag{4.6}$$

Note that the $J$ columns of resulting $H$ correspond to the SFR rotated in directions $\{\mathbf{v}_j\}_{j=1}^J$ and discretized at points $\{\mathbf{u}_k\}_{k=1}^K$. Finally, we define $\Phi$ to be a $K \times (J+1)$ matrix obtained from $H$ through addition of an extra column of ones [67]

$$\Phi = [H \quad \mathbf{1}], \tag{4.7}$$

[1] To estimate these parameters, it is standard to fit the DTI model of (4.1) to HARDI signals corresponding to, e.g., corpus callosum. As the latter is predominantly composed of single (commisural) fibre bundles, its related signals can provide a reliable estimate of the SFR, subject to appropriate averaging [33].

where $\mathbf{1} = [1, 1, ..., 1]^T \in \mathbb{R}^K$.

Now, let $\mathbf{r}_i$ be the coordinate of an arbitrary voxel within $\Omega$, with $i = 1, 2, \ldots, I$, and let $f^i \in \mathbb{R}^{J+1}$ be a column vector defined as $f^i = [f'_a(\mathbf{v}_1|\mathbf{r}_i), f'_a(\mathbf{v}_2|\mathbf{r}_i), \ldots, f'_a(\mathbf{v}_J|\mathbf{r}_i), f'_{\text{iso}}(\mathbf{r}_i)]^T$. Then, in the absence of measurement noises and disregarding the effect of discretization, the SD model (4.4) along with (4.5) suggest that $s^i = \Phi f^i$, for all $i = 1, 2, \ldots, I$. It is, therefore, convenient to agglomerate all the above model equations into a single one that reads [67]

$$s = \Phi f, \tag{4.8}$$

where $f$ is a $(J + 1) \times I$ matrix, with its columns defined by $f^i$, with $i = 1, 2, \ldots, I$. Note that the last row $f_{J+1}$ of $f$ is equal to a row-stacked version of the IDM, whilst the first $J$ rows $f_1, \ldots, f_J$ of $f$ can be viewed as row-stacked versions of the images obtained by restricting the fODF $f_a(\mathbf{u}|\mathbf{r})$ to the directions $\mathbf{v}_1, \ldots, \mathbf{v}_J$, respectively.

### 4.2.2 Estimation Framework

For obvious reasons, recovering a useful estimate of $f$ based on model (4.8) alone is a futile exercise. According to the formalism of Bayesian estimation, to render the reconstruction unique and stable, the model equation needs to be augmented with reasonable a priori assumptions on the nature of $f$ in (4.8). One of such assumptions, which has been proved to be particularly useful for reconstruction of fODFs is sparsity, discussed in Section 3.1.2. Indeed, the anatomical organization of white matter suggests the number of axonal fascicles running through any given voxel $\mathbf{r}_i$ is likely to be relatively small. This fact, in turn, implies that the vector $[f'_a(\mathbf{v}_1|\mathbf{r}_i), f'_a(\mathbf{v}_2|\mathbf{r}_i), \ldots, f'_a(\mathbf{v}_J|\mathbf{r}_i)]^T$ could be reasonably expected to have a relatively small number of significant components, with the rest of its entries distributed in a close proximity of zero. Such a behaviour of $f_a(\mathbf{u}|\mathbf{r})$ can be modelled in a number of different ways [40, 46, 44]. In this thesis, we take advantage of the standard method of recovering a sparse vector through minimization of its $\ell_1$-norm [40, 49].

Before proceeding to the next step, we note that the SD model of (4.8) admits an alternative interpretation, according to which every $s^i$ is approximated by a linear combination of the columns of $\Phi$. As the last (constant) column of $\Phi$ has distinctly different morphological properties as compared to the other columns of the matrix, it seems reasonable to minimize the $\ell_1$-norm of the entire vector $f^i$ (rather than only of its part associated with $f'_a(\mathbf{u}|\mathbf{r}_i)$). In this case, the parsimonious nature of $\ell_1$ minimization will force the optimal solution to be dominated by either an IDM or an fODF component, while permitting them both only when there is clear evidence of their concurrent existence. It deserves noting

that a similar principle has been used in morphological component analysis to decompose a signal of interest into morphologically distinct components [77, 78].

To avoid unnecessary complications in notations, in the derivations that follow, we use $\|\cdot\|_2$ and $\|\cdot\|_1$ to denote the $\ell_2$- and $\ell_1$-norms of vectors as well as the analogous "entry-wise" norms of matrices. Thus, for example, the $\ell_2$-norm of a HARDI signal $s$ can be expressed in two ways as $\|s\|_2^2 = \sum_{i=1}^{I} \|s^i\|_2^2 = \sum_{k=1}^{K} \|s_k\|_2^2$, while the $\ell_1$-norm of $f$ can be expressed as $\|f\|_1 = \sum_{i=1}^{I} \|f^i\|_1 = \sum_{j=1}^{J+1} \|f_j\|_1$. With this notation at hand, we formulate the problem of finding an optimal $f^*$ as

$$\begin{aligned} f^* &= \arg\min_f \|f\|_1 \\ \text{subject to } &\|\Phi f - s\|_2^2 \leq \epsilon,\ f \geq 0, \end{aligned} \tag{4.9}$$

where $\epsilon$ controls the size of measurement and model errors, whereas the point-wise inequality constraint is added to assure that both the IDM and fODFs are non-negative quantities. Note that the problem in (4.9) is analogous to the one described in [44], apart from the fact that (4.9) applies to a whole set of HARDI data, rather than to a single voxel.

To facilitate numerical solution, the problem (4.9) is usually reformulated in its equivalent, unconstrained (Lagrangian) form as given by [67]

$$f^* = \arg\min_f \left\{ \frac{1}{2}\|\Phi f - s\|_2^2 + \lambda \|f\|_1 + \varphi_{\geq}(f) \right\}, \tag{4.10}$$

where $\lambda > 0$ is a user-controlled regularization parameter and $\varphi_{\geq}$ denotes the indicator function of the positive orthant. Specifically, $\varphi_{\geq}(f) = 0$, if all entries of $f$ are non-negative, and $\varphi_{\geq}(f) = +\infty$, otherwise. The problem (4.10) can be solved using a variety of methods of non-smooth optimization (e.g., [79]).

Unfortunately, the solution of (4.10) could be only suboptimal, as it completely disregards any spatial-domain dependencies between the values of $f$. To palliate this deficiency, we first note that each row of $f$ can be considered to be a discrete image defined over the spatial lattice $\{\mathbf{r}_i\}_{i=1}^{I} \in \Omega$ and stacked into a row vector. Particularly, in this interpretation, the first $J$ rows $\{f_j\}_{j=1}^{J}$ can be viewed as restrictions of the fODF to directions $\{\mathbf{v}_j\}_{j=1}^{J}$, while $f_{J+1}$ represents the corresponding IDM. Naturally, the images $\{f_j\}_{j=1}^{J}$ and $f_{J+1}$ have different statistical properties, and therefore they should be regularized in different ways. Thus, to spatially regularize the fODF $f_a(\mathbf{u}|\mathbf{r})$ we adopt the fibre continuity approach of [50], which requires the directional derivative of $f_j$ along $\mathbf{v}_j$ to be relatively small in value. Formally, let $\nabla_d : \mathbb{R}^I \to \mathbb{R}^I$, with $d = 1, 2, 3$, denote the operators of spatial differencing in the direction of $x$, $y$, and $z$ coordinates, respectively. Then, for each $j = 1, \ldots, J$, one

can then assemble a $3 \times I$ matrix $Df_j$, with its rows defined by the partial differences of $f_j$, *viz.* $Df_j = [\nabla_1 f_j^T \ \nabla_2 f_j^T \ \nabla_3 f_j^T]^T \in \mathbb{R}^{3\times I}$. Subsequently, with $\mathbf{v}_j$ being a column vector, the directional derivative of $f_j$ along $\mathbf{v}_j$ is conveniently given by $\mathbf{v}_j^T Df_j$, in which case the approach of [50] calls for minimizing $\sum_{j=1}^{J} \|\mathbf{v}_j^T Df_j\|_2^2$.

It should be noted that the above approach cannot be extended to the IDM $f_{J+1}$, since the latter is devoid of directional continuity [50]. Yet, for the pure sake of harmonizing the notations, we replace the original minimization of $\sum_{j=1}^{J} \|\mathbf{v}_j^T Df_j\|_2^2$ by minimizing $\sum_{j=1}^{J+1}(1-\delta_{j,J+1})\|\mathbf{v}_j^T Df_j\|_2^2$, where $\delta_{t,r}$ stands for the Kronecker symbol which obeys $\delta_{t,r} = 1$, if $t = r$, while $\delta_{t,r} = 0$, otherwise. (Note that, since the last component in the above summation is multiplied by a zero weight, the choice of $\mathbf{v}_{J+1}$ is immaterial and it has no effect on the reconstruction procedure as shown later.) Subsequently, the optimal solution can still be defined as a global minimizer over $f$, which is now given by

$$f^* = \arg\min_f \left\{ \frac{1}{2}\|\Phi f - s\|_2^2 + \lambda\|f\|_1 + \mu\|f\|_a^2 + \varphi_{\geq}(f) \right\}, \tag{4.11}$$

where

$$\|f\|_a^2 = \sum_{j=1}^{J+1}(1-\delta_{j,J+1})\|\mathbf{v}_j^T Df_j\|_2^2, \tag{4.12}$$

and $\mu > 0$ is another regularization constant.

Finally, the problem of regularizing the IDM $f_{J+1}$ must not be overseen as well. Since one can reasonably expect $f_{J+1}$ to vary smoothly throughout the brain with the exception of abrupt changes between white matter, grey matter, CSF, as well as the regions of possible brain pathologies, it seems justified to model $f_{J+1}$ as a function of bounded variation (BV) [80]. In particular, the BV model enforces the assumption on $f_{J+1}$ to have a relatively small value of its total variation (TV) seminorm $\|f_{J+1}\|_{TV}$ which can be defined as follows. Let $\nabla_1 f_{J+1}[i]$, $\nabla_2 f_{J+1}[i]$, and $\nabla_3 f_{J+1}[i]$ denote the $i$-th elements of vectors $\nabla_1 f_{J+1}$, $\nabla_2 f_{J+1}$, and $\nabla_3 f_{J+1}$, respectively. Then, the TV seminorm of $f_{J+1}$ can be defined in a standard way as

$$\|f_{J+1}\|_{TV} = \sum_{i=1}^{I} \Big[ \sum_{d=1}^{3} |\nabla_d f_{J+1}[i]|^2 \Big]^{1/2}. \tag{4.13}$$

It is definitely possible to apply the definition of TV to images $f_j$, with $j = 1, \ldots, J$ as well. However, minimizing these norms would likely mislead the estimation process, since $f_j$ may not be assumed to be piecewise smooth, in general. Still, to balance the notations,

we define the TV seminorm of $f$ according to

$$\|f\|_{TV} = \sum_{j=1}^{J+1} \delta_{j,J+1} \|f_j\|_{TV}, \tag{4.14}$$

which leads to the optimal solution of the form

$$f^* = \arg\min_f \left\{ \frac{1}{2}\|\Phi f - s\|_2^2 + \lambda\|f\|_1 + \mu\|f\|_a^2 + \nu\|f\|_{TV} + \varphi_{\geq}(f) \right\}, \tag{4.15}$$

with $\nu > 0$ being an additional regularization constant which controls the piecewise smooth behaviour of the IDM. We admit that automatically determining an optimal value of $\nu$, as well as those of $\lambda$ and $\mu$, is a difficult problem, which extends well beyond the scope of the this thesis. It was observed in practice, however, that finding acceptable values of these parameters by wonted trials-and-errors is a much less arduous task, as it might seem at the first glance. To highlight the applied regularization terms in (4.15), the SD method proposed here will be referred below to as SCSD (standing for spatially constrained sparse deconvolution).

The solution of (4.15) entails minimizing a non-smooth cost function, which effectively rules out the use of gradient-based methods of numerical optimization. Moreover, the composite nature of the cost makes it difficult to devise an efficient optimization approach which would perform the minimization directly with respect to $f$. To overcome these difficulties, the next section introduces a particularly simple solution using the alternating directions method of multipliers (ADMM) [81]. Apart from breaking down the optimization in (4.15) into a sequence of simple and closed-form solutions, the method offers a straightforward approach to splitting the computations between multiple computing cores/units, which is a significant advantage considering the relatively large dimensionality of HARDI data.

## 4.3 Numerical Solution

To simplify the solution of (4.15), we introduce two auxiliary variables $u$ and $v$, and replace the original optimization problem by an equivalent, equality-constrained one. Specifically,

$$\min_{f,u,v} \left\{ \frac{1}{2}\|\Phi f - s\|_2^2 + \lambda\|u\|_1 + \mu\|v\|_a^2 + \nu\|v\|_{TV} + \varphi_{\geq}(u) \right\}, \tag{4.16}$$
$$\text{subject to } f = u,\ f = v$$

Note that in (4.16) the minimization is carried out with respect to three variables, namely $f$, $u$, and $v$. A standard approach to solving such equality-constrained problems is based of the use of augmented Lagrangian methods [82]. Particularly, for the case at hand, this approach amounts to the following iterations (starting from some intitial values of Lagrange multipliers $p_u^0$ and $p_v^0$, e.g., $p_u^0 = p_v^0 = 0$).

$$\begin{aligned}(f^{(t+1)}, u^{(t+1)}, v^{(t+1)}) = \arg\min_{f,u,v} \Big\{ &\frac{1}{2}\|\Phi f - s\|_2^2 + \lambda\|u\|_1 + \mu\|v\|_a^2 + \nu\|v\|_{TV} + \\ &+\varphi_{\geq}(u) + \frac{\delta_u}{2}\|f - u + p_u^{(t)}\|_2^2 + \frac{\delta_v}{2}\|f - v + p_v^{(t)}\|_2^2 \Big\} \qquad (4.17)\\ p_u^{(t+1)} = p_u^{(t)} + f^{(t+1)} - u^{(t+1)} \\ p_v^{(t+1)} = p_v^{(t)} + f^{(t+1)} - v^{(t+1)}\end{aligned}$$

where $t$ stands for an iteration index, while $\delta_u > 0$ and $\delta_v > 0$ are some positive constants[2].

Finally, in ADMM, the concurrent minimization with respect to $f$, $u$, and $v$ is replaced by sequential minimization with respect to $f$, $u$, and $v$ independently. In this case, starting from some $f^{(0)} = u^{(0)} = v^{(0)}$, the update proceeds according to:

$$\textbf{Step1}: f^{(t+1)} = \arg\min_{f} \Big\{ \frac{1}{2}\|\Phi f - s\|_2^2 + \frac{\delta_u}{2}\|f - u^{(t)} + p_u^{(t)}\|_2^2 + \frac{\delta_v}{2}\|f - v^{(t)} + p_v^{(t)}\|_2^2 \Big\} \qquad (4.18)$$

$$\textbf{Step2}: u^{(t+1)} = \arg\min_{u} \Big\{ \frac{\delta_u}{2}\|f^{(t+1)} - u + p_u^{(t)}\|_2^2 + \lambda\|u\|_1 + \varphi_{\geq}(u) \Big\} \qquad (4.19)$$

$$\textbf{Step3}: v^{(t+1)} = \arg\min_{v} \Big\{ \frac{\delta_v}{2}\|f^{(t+1)} - v + p_v^{(t)}\|_2^2 + \mu\|v\|_a^2 + \nu\|v\|_{TV} \Big\}, \qquad (4.20)$$

followed by the "dual" step of updating the Lagrange multipliers $p_u$ and $p_v$, as specified in (4.17). Although the ADMM approach has effectively supplanted a single minimization problem (4.15) by three minimization problems, the latter admit much simpler solutions as detailed below.

### 4.3.1 Solution to Step 1

The optimization problem in (4.18) is a classical least-square (LS) problem which admits a closed-form solution as given by

$$f^{(t+1)} = \big(\Phi^T\Phi + (\delta_u + \delta_v)I_{J+1}\big)^{-1}\big(\Phi^T s + \delta_u(u^{(t)} - p_u^{(t)}) + \delta_v(v^{(t)} - p_v^{(t)})\big), \qquad (4.21)$$

[2]Note that the algorithm is guaranteed to converge for any positive $\delta_u$ and $\delta_v$. In the present work we use $\delta_u = \delta_v = 0.5$.

where $I_{J+1}$ stands for a $(J+1)\times(J+1)$ identity matrix. To facilitate the computations, the inverse matrix $R \triangleq \left(\Phi^T\Phi + (\delta_u+\delta_v)I_{J+1}\right)^{-1}$ can be precomputed and stored before the reconstruction procedure is initiated.

### 4.3.2 Solution to Step 2

The optimization problem in (4.19) can be equivalently rewritten as

$$u^{(t+1)} = \arg\min_u \left\{\frac{1}{2}\|u-(f^{(t+1)}+p_u^{(t)})\|_2^2 + \frac{\lambda}{\delta_u}\|u\|_1 + \varphi_{\geq}(u)\right\}. \tag{4.22}$$

In the absence of the indicator function $\varphi_{\geq}$, the solution to the above problem would be given by soft thresholding, *viz.* $u^{(t+1)} = \mathcal{S}_{\lambda/\delta_u}\{f^{(t+1)}+p_u^{(t)}\}$, with $\mathcal{S}_\tau(x) = \text{sign}(x)(|x|-\tau)_+$ (where the subscript + denotes the operation of keeping the positive part of the argument). As unexpected as it might sound, however, incorporating the indicator function actually simplifies the solution by replacing the operator $\mathcal{S}_\tau(x)$ with its positively rectified version $\mathcal{S}_\tau^+(x) = (x-\tau)_+$. Consequently, the solution to (4.19) is given by

$$u^{(t+1)} = \mathcal{S}^+_{\lambda/\delta_u}\{f^{(t+1)}+p_u^{(t)}\} = \left(f^{(t+1)}+p_u^{(t)}-\lambda/\delta_u\right)_+. \tag{4.23}$$

### 4.3.3 Solution to Step 3

To derive an update equation for $v^{(t)}$, it is convenient to rewrite the optimization problem in (4.20) in a slightly different form as given by

$$v^{(t+1)} = \arg\min_v \left\{\frac{1}{2}\|v-(f^{(t+1)}+p_v^{(t)})\|_2^2 + \frac{\mu}{\delta_v}\|v\|_a^2 + \frac{\nu}{\delta_v}\|v\|_{TV}\right\} \tag{4.24}$$

Furthermore, since in our (simplified) notations $\|v\|_2^2 = \sum_{j=1}^{J+1}\|v_j\|_2^2$, then using the definitions of $\|v\|_a^2$ and $\|v\|_{TV}$ in (4.12) and (4.14), correspondingly, one can rewrite (4.24) in an equivalent form as

$$\begin{aligned} v^{(t+1)} = \arg\min_v \Bigg\{ \sum_{j=1}^{J+1} \Bigg(&\frac{1}{2}\|v_j-(f^{(t+1)}+p_v^{(t)})_j\|_2^2 + \\ &+ \frac{\mu}{\delta_v}(1-\delta_{j,J+1})\|\mathbf{v}_j^T D v_j\|_2^2 + \frac{\nu}{\delta_v}\delta_{j,J+1}\|v_j\|_{TV}\Bigg)\Bigg\}, \end{aligned} \tag{4.25}$$

where $(f^{(t+1)}+p_v^{(t)})_j$ denotes the $j$-th row of matrix $f^{(t+1)}+p_v^{(t)}$.

A closer look at (4.25) reveals that its cost functional consists of $J+1$ positive-valued terms, each of which depends on the variables $v_j$, $j = 1, 2, \ldots, J+1$, independently. Hence, the minimization over $v$ can be replaced by minimizations of the $J+1$ summands in (4.25) with respect to their respective variables (i.e., $v_j$). In particular, for $j = 1, \ldots, J$, the resulting minimization problems are given by

$$v_j^{(t+1)} = \arg\min_{v_j} \left\{ \frac{1}{2}\|v_j - (f^{(t+1)} + p_v^{(t)})_j\|_2^2 + \frac{\mu}{\delta_v}\|\mathbf{v}_j^T D v_j\|_2^2 \right\}. \tag{4.26}$$

The above is a simple LS problem, which can be solved using spectral methods (i.e., by means of linear filtering) as detailed in the Appendix. Note that the cost function in (4.26) does not contain a TV term due to mutual exclusivity of the weights in definitions of $\|v\|_a^2$ and $\|v\|_{TV}$. For the same reason, optimization over $v_{J+1}$ does not contain a fibre continuity term, resulting in

$$v_{J+1}^{(t+1)} = \arg\min_{v_{J+1}} \left\{ \left( \frac{1}{2}\|v_{J+1} - (f^{(t+1)} + p_v^{(t)})_{J+1}\|_2^2 + \frac{\nu}{\delta_v}\|v_{J+1}\|_{TV} \right\}. \tag{4.27}$$

Note that the problem in (4.27) is a classical TV-regularized denoising problem [80], which can be efficiently solved by a variety of different algorithms. In the current work, we used the semi-implicit, fixed-point approach of [83] due to its impressive numerical stability and fast convergence.

To summarise the iterative steps of solving optimization problem (4.16), a pseudocode is given as the Fig. 4.1.

Proposed Algorithm: SCSD

**Data**: diffusion weighted signal $s$, convolution matrix $\Phi$, $\lambda$, $\mu$, $\nu$, $\delta_u$ $\delta_v$, number of iterations $NIT$, $tol$

**Result**: $f$

initialization;

$u^0 \leftarrow 0_{(J+1)\times I}$

$v^0 \leftarrow 0_{(J+1)\times I}$

$p_u^0 \leftarrow 0_{(J+1)\times I}$

$p_v^0 \leftarrow 0_{(J+1)\times I}$

$t \leftarrow 0$

**while** $error > tol$ *and* $t < NIT$ **do**

$f^{(t+1)} \leftarrow \left(\Phi^T\Phi + (\delta_u + \delta_v)I_{J+1}\right)^{-1}\left(\Phi^T s + \delta_u(u^{(t)} - p_u^{(t)}) + \delta_v(v^{(t)} - p_v^{(t)})\right)$,

$u^{(t+1)} \leftarrow \left(f^{(t+1)} + p_u^{(t)} - \lambda/\delta_u\right)_+$,

$v_j^{(t+1)} \leftarrow \arg\min_{v_j}\left\{\frac{1}{2}\|v_j - (f^{(t+1)} + p_v^{(t)})_j\|_2^2 + \frac{\mu}{\delta_v}\|\mathbf{v}_j^T D v_j\|_2^2\right\}^*$,

$v_{J+1}^{(t+1)} \leftarrow \arg\min_{v_{J+1}}\left\{\left(\frac{1}{2}\|v_{J+1} - (f^{(t+1)} + p_v^{(t)})_{J+1}\|_2^2 + \frac{\nu}{\delta_v}\|v_{J+1}\|_{TV}\right\}^{**}\right.$,

$p_u^{(t+1)} \leftarrow p_u^{(t)} + f^{(t+1)} - u^{(t+1)}$,

$p_v^{(t+1)} \leftarrow p_v^{(t)} + f^{(t+1)} - v^{(t+1)}$.

$error \leftarrow \frac{1}{2}\|\Phi f - s\|_2^2 + \lambda\|u\|_1 + \mu\|v\|_a^2 + \nu\|v\|_{TV}$

$cntr++$

**end**

* A simple LS problem, which can be solved using spectral methods as in the Appendix.

** A TV-regularized denoising problem can be solved by different algorithms, e.g.[83].

Figure 4.1: Pseudocode for iteratively solving the optimization problem in SCSD.

# Chapter 5

# Experiments and Results

The introduction of in silico and in vivo data in this chapter has been presented in a conference paper [67], which also gave the same definitions of two comparison metrics, TP and AAE, reused here. Meanwhile parts of illustration figures and experiment results shown in this chapter are reprinted from that conference paper.

## 5.1 Materials and Methods

### 5.1.1 Sources of Data

The performance of the proposed and reference methods has been assessed using both computer-simulated and in vivo data. Specifically, the simulated data were generated based on a standard Gaussian mixture model [25] in (4.2) with $\lambda_{\parallel} = 17 \cdot 10^{-4}$ mm$^2$/s and $\lambda_{\perp} = 3 \cdot 10^{-4}$ mm$^2$/s. The spatial dimensions of the data were set to be equal to $16 \times 16 \times 12$, while the directions of diffusion encoding were defined by the 2nd order tessellation of icosahedron restricted to the northern hemisphere of $\mathbb{S}^2$ (thus resulting in $K = 81$ sampling points). The data were designed so as to emulate a crossing of two cylindrically symmetric "fibres" of 8 voxels in diameter, with the crossing angle $\alpha$ in the range $\alpha \in [30°, 90°]$ with step size 5°. Subplots A and D of Fig. 5.1 depict the theoretical ODFs corresponding to the middle "layer" of two signal arrays which have been synthesized for $\alpha = 60°$ and $\alpha = 45°$, respectively. Additionally, to allow investigation of the effect of isotropic diffusion on the accuracy of SD based reconstruction, the simulated signals have also been combined with a constant (i.e., isotropic) component of a variable magnitude.

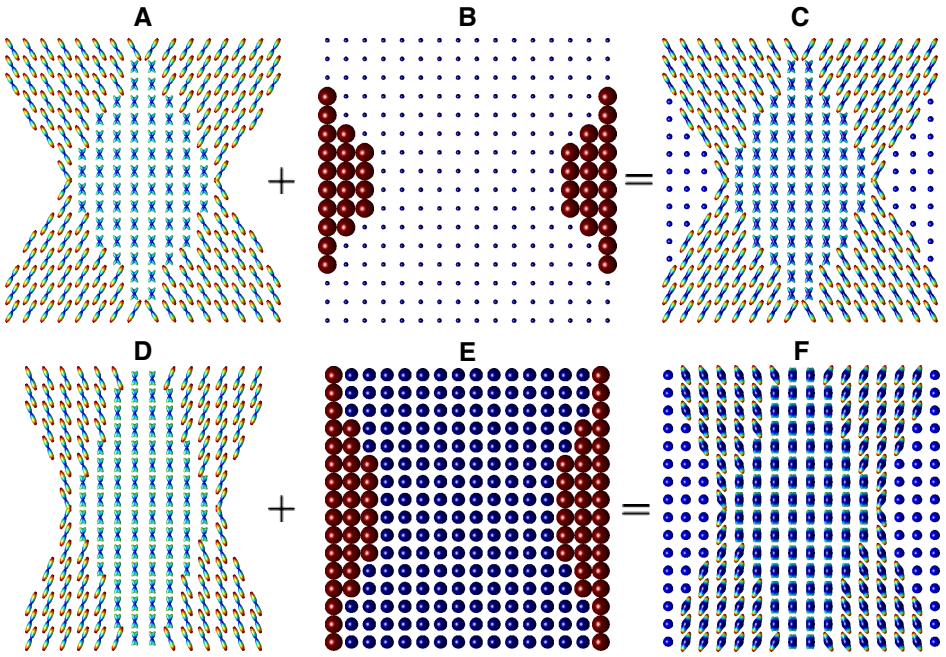


Figure 5.1: (Subplots A-C, left-to-right) The anisotropic, isotropic, and the combined field of ODFs associated with in silico HARDI signals generated with $b = 3000$ s/mm$^2$, $\alpha = 60°$, and $p_{iso} = 0.25$; (Subplots D-F) Same as above, only with $\alpha = 45°$ and $p_{iso} = 0.75$. Reprinted from [67, p.5].

The diffusivity of the isotropic component was set to be equal to $8 \cdot 10^{-4}$ mm$^2$/s, whereas its partial volume fractions outside and inside of the "fibres" were set to 1 and $p_{iso}$, respectively, with $p_{iso} \in \{0, 0.25, 0.5, 0.75\}$. Two examples of the isotropic component are shown in Subplots B and E of Fig. 5.1 for the case of $p_{iso} = 0.25$ and $p_{iso} = 0.75$, correspondingly. (Note that, in the above subplots, the absolute size of the glyphs has been optimized for visualization, and thus it does not represent the actual values of the isotropic ODFs.) Finally, Subplot C of Fig. 5.1 depicts combined ODFs obtained as a result of the summation of the ODFs shown in Subplots A and B, while the result of the summation of the ODFs in Subplots D and E is displayed in Subplot F of the same figure.

The data simulation was repeated for two different values of $b$, *viz.* $b = 1000$ s/mm$^2$ and $b = 3000$ s/mm$^2$, with the corresponding SNRs of Rician noise[1] being 20 and 7. It

[1]In the case at hand, SNR was defined to be a ratio of the mean amplitude of a noise-free HARDI signal to the standard deviation of the Gaussian noise contaminating the complex MR readout.

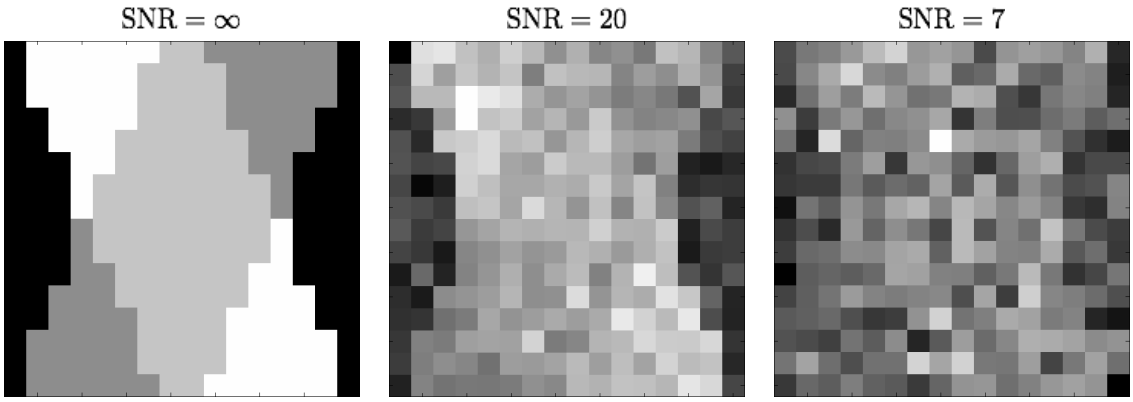


Figure 5.2: Example of a noise-free "axial slice" of the simulated HARDI data (left) along with its noise contaminated versions for SNR = 20 (middle) and SNR =7 (right). Reprinted from [67, p.6].

should be noted that the above choice of $b$-values and their related SNRs is by no means arbitrary, but intended to imitate a real-life situation, in which increasing the value of $b$ comes at the price of a substantially reduced SNR. (The effect of Rician noise is exemplified in Fig. 5.2 where a noise-free "axial slice" of the simulated HARDI signal (left) is shown along with its noise contaminated versions for SNR = 20 (middle) and SNR=7 (right).) In such a case, the adverse effect of measurement noises could effectively counterpoise the gain in signal bandwidth which comes with exploiting higher $b$-values. In the context of SD, it is therefore important to understand the trade-off between the values of $b$ and SNR under various regularization schemes.

Besides the computer simulated data, we also did experiments with in vivo data which were provided by the Centre for Functional and Metabolic Mapping (CFMM) at Western's Robarts Research Institute. Real-life HARDI data were acquired by means of a 3T Siemens TrioTim MRI scanner. The diffusion-encoded scans were collected from a 45 y.o. healthy volunteer at 2 mm isotropic resolution with: $b \in \{1000, 3000\}$ s/mm$^2$, TR=6300 ms, and TE = 85 ms. The transversal dimensions of the scans were equal to 128×128 pixels, with a total of 50 axial slices used. The diffusion-encoding directions $\{\mathbf{u}_k\}_{k=1}^{K}$ (with $K = 64$) were defined using the method of generalized spirals [84], which provides a closed-form solution to the problem of quasi-uniform sampling of $\mathbb{S}^2$.

### 5.1.2 Reference Methods

In this work, we compared the performance of the proposed method against that of a number of alternative approaches. The first of these approaches was the positively constrained LS algorithm of [35], which will be referred below to as CSD. Note that this method disregards the presence of isotropic diffusion, and its solutions can be produced by solving (4.15) upon replacing $\Phi$ by $H$ and setting $\lambda = \mu = \nu = 0$. In the case when $\lambda > 0$, CSD transforms into the method of [40], which was the second reference method (referred below to as Min-$\mathcal{L}_1$) used in our comparative study. Note that, just like CSD, Min-$\mathcal{L}_1$ disregards both the presence of isotropic diffusion and the spatial-domain regularity of estimated fODFs. While still ignoring the effect of isotropic diffusion, the method of [50] takes advantage of an original way to constrain the spatial behaviour of reconstructed fODFs using a fibre continuity model. This method (which we refer to below as CSD-FC) can be also described by the minimization problem in (4.15) upon replacing $\Phi$ by $H$ and setting $\lambda = \nu = 0$. Further, setting $\mu = 0$ in (4.15) would result in the SD algorithm recently proposed in [67]. This method explicitly accounts for the presence of isotropic diffusion, while constraining the spatial regularity of its related IDM. However, the method of [67] (referred below to as Min-TV-$\mathcal{L}_1$) does not enforce the spatial continuity of fibre orientations, as it is done in the case of the proposed algorithm. In [42], the authors proposed an SD approach based on the famed Lucy-Richardson deconvolution procedure (referred to as dRL). Although disregarding the spatial-domain regularity of the estimated fODFs, dRL had been particularly designed to suppress the influence of isotropic diffusion on the deconvolution results.

### 5.1.3 Comparison Metrics

To quantitatively assess the results produced by different SD routines, a total of four comparison metrics have been employed. The first of these metrics is average angular error (AAE) defined as [67]

$$\mathrm{AAE} = \frac{180^\circ}{\pi}\mathcal{E}\{\arccos(\mathbf{v}_0 \cdot \tilde{\mathbf{v}})\}, \tag{5.1}$$

where $\mathbf{v}_0$ and $\tilde{\mathbf{v}}$ stand for an original fibre orientation and its estimate respectively and $\mathcal{E}$ denotes the operator of statistical expectation (which is commonly approximated by a sample mean in practice). It should be noted that a standard way to infer the value(s) of $\tilde{\mathbf{v}}$ from an estimated fODF is through finding the local maxima of the latter. Alternatively, one could use the estimated fODF to fit to it a model of the form $\sum_{l=1}^{M} \exp(\mathbf{u}^T A_l \mathbf{u})$ (with $M$ denoting the number of fibres and $A_l$ being a symmetric $3\times3$ matrix) and then associate $\tilde{\mathbf{v}}$ with the direction of the major eigenvector of $A_l$. When the number of fibres $M$ is known

as *a priori* (as it is the case with computer-simulated studies), the second approach tends to produce more accurate results. For this reason, it was exploited in this thesis.

Additional metrics used in our comparative analysis were true positive (TP) and false positive (FP) rates of detection of the correct number of fibres at each voxel. Specifically, TP can be defined to be a percentage of data voxels, in which the number of local maxima of an estimated fODF coincided with the true number of simulated fibres. To render the computation of TP more robust towards overestimation errors, at each voxel the local maxima have been subjected to hard thresholding, set at the level of 20% of their peak value. The FP metric, on the other hand, was particularly designed to assess the extent of overestimation of the true number of fibres [50, 42]. Formally, FP can be defined as

$$\mathrm{FP} = \mathcal{E}\{(\tilde{M} - M_0)_+\}, \tag{5.2}$$

where $M_0$ and $\tilde{M}$ denote the true number of fibres and the number of local maxima in reconstructed fODFs. Note that, similarly to the case of TP, before evaluating FP, all the local maxima had been subjected to the same hard thresholding procedure.

Finally, one of the principal outputs of the proposed method is an estimate of the IDM, which is a scalar-valued function of the spatial coordinate. In the case at hand, the quality of reconstructed IDMs can be assessed qualitatively based on the notion of image contrast. The latter can be defined as follows. First, we partition the entire image domain $\Omega$ into two subdomains, $\Omega_{in}$ and $\Omega_{out}$ (with $\Omega = \Omega_{in} \cup \Omega_{out}$), which encompass the regions occupied by the simulated fibres and purely isotropic diffusion (i.e., no fibres present), respectively. Also, we denote by $f_{iso}(\Omega_{in})$ and $f_{iso}(\Omega_{out})$ the restrictions of $f_{iso}$ to the corresponding subdomains. Then, letting $\mu_{in}$ and $\mu_{out}$ (resp. $\sigma_{in}$ and $\sigma_{out}$) denote the mean values (resp. the standard deviations) of $f_{iso}(\Omega_{in})$ and $f_{iso}(\Omega_{out})$, correspondingly, a measure of contrast $C$ can be defined as given by

$$C = 2\,\frac{|\mu_{in} - \mu_{out}|}{\sigma_{in} + \sigma_{out}}. \tag{5.3}$$

It is worthwhile noting that, the (real) IDMs used in the experimental study are piecewise constant "images", in which case $\sigma_{in} = \sigma_{out} = 0$, resulting in $C = \infty$. Thus, in the case of estimated IDMs, it is reasonable to assume that higher values of $C$ represent more accurate estimates of the original IDMs.

As a closing remark we note that, among all the SD methods under comparison, only Min-TV-$\mathcal{L}_1$ and SCSD are capable (by design) of computing the estimates of IDM, whereas neither of the remaining methods is endowed with the same capability. However, both CSD and Min-$\mathcal{L}_1$, as well as CSD-FC, have been proposed as means for estimation of the fODF $f_a$, in which case the remainder $H\tilde{f}_a - s$ (with $\tilde{f}_a$ being an estimate of $f_a$) is likely to

represent estimation errors along with a contribution from $f_{iso}$. Furthermore, assuming the estimation errors to have a zero mean value, a reasonable estimate $\tilde{f}_{iso}$ of $f_{iso}$ can be obtained by averaging the remainder $H\tilde{f}_a - s$ over the spherical coordinate. Formally, one can define $\tilde{f}_{iso}$ to be

$$\tilde{f}_{iso} \simeq \frac{A}{K}\,\mathbf{1}^T\big(H\tilde{f}_a - s\big), \tag{5.4}$$

where $\mathbf{1}$ denotes a $K$-dimensional (column) vector of ones and $A > 0$ is a proportionality constant that could be set to 1, since it does not have any effect on the value of $C$ in (5.3). We also note that, by its design, the dRL algorithm of [42] aims to recover an FOF (rather than an fODF) as given by (4.5), thereby making the estimation of IDMs according to (5.4) inadequate. Consequently, quantitative comparisons in terms of $C$ have been performed for all the methods under consideration, except for dRL.

## 5.2 Results

### 5.2.1 Computer Simulations

The left and right columns of subplots in Fig. 5.3 show the values of AAE as a function of crossing angle $\alpha$, which have been obtained with the proposed and reference methods for $b = 1000$ s/mm$^2$ and $b = 3000$ s/mm$^2$, respectively. The regularization parameters for Min-$\mathcal{L}_1$, CSD-FC, Min-TV-$\mathcal{L}_1$, and SCSD were set to optimize the overall performance of the algorithms, namely: 1) $\lambda = 0.01$ for Min-$\mathcal{L}_1$, 2) $\mu = 0.01$ for CSD-FC, 3) $\lambda = 0.07$, $\nu = 0.01$ for Min-TV-$\mathcal{L}_1$, and 4) $\lambda = 0.03$, $\nu = 0.01$, $\mu = 0.4$ for SCSD. The dRL method has been reproduced following its description in [42]. One can see that all the error curves, with the exception of that of dRL, exhibit the expected behaviour where AAE decreases with an increase in $\alpha$. Moreover, despite the substantially worse noise conditions for $b = 3000$ s/mm$^2$, all the tested methods (again, with the exception of dRL) demonstrate better performance for $b = 3000$ s/mm$^2$, as compared to the case of $b = 1000$ s/mm$^2$. In all the cases, however, the proposed SCSD method shows considerably better performance in comparison to the alternative solutions, with the "second best" results produced by CSD-FC for lower values of $p_{iso}$ and by Min-TV-$\mathcal{L}_1$ for higher values of $p_{iso}$.

Before proceeding any further, one additional comment is in order regarding the behaviour of the AAE curves obtained with dRL. Specifically, one can see that, for $b = 1000$ s/mm$^2$, the AAE is minimized for smaller values of $\alpha$, which is rather a counter-intuitive result. To understand why this happens, it is instructive to examine the behaviour of the TP curves produced by dRL (see Fig. 5.4). Specifically, one can see that for $b = 1000$ s/mm$^2$

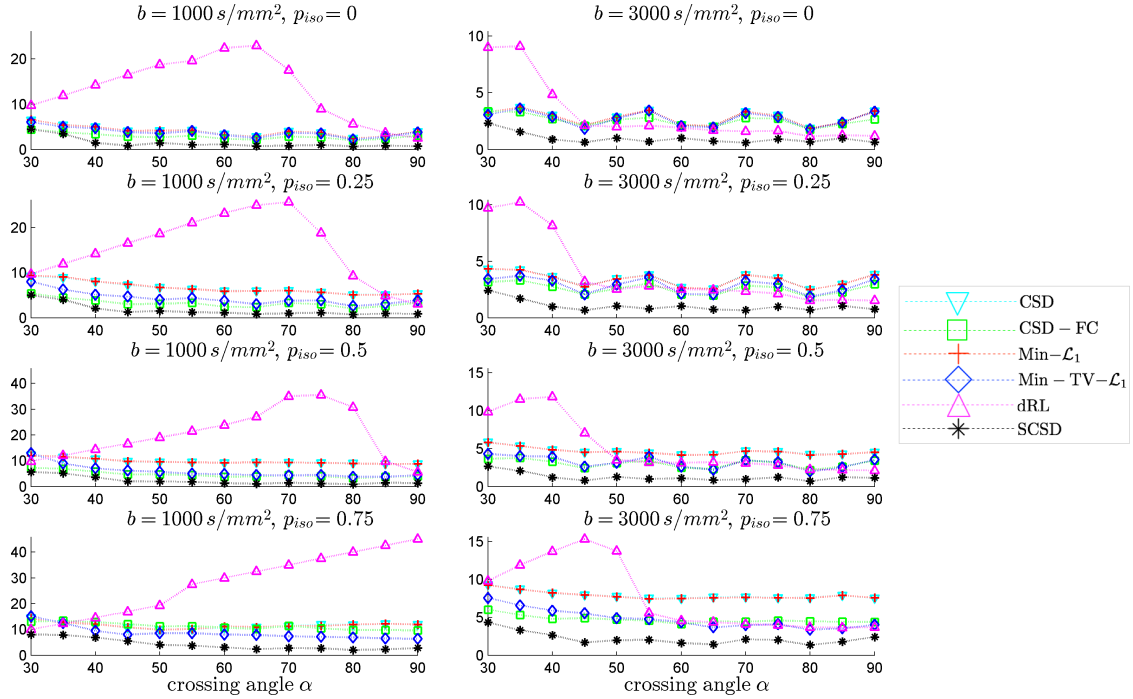


Figure 5.3: (Left columns of subplots) AAE produced by the tested methods for different $\alpha$, $p_{iso}$, and $b = 1000$ s/mm$^2$; (Right columns of subplots) AAE produced by the tested methods for different $\alpha$, $p_{iso}$, and $b = 3000$ s/mm$^2$. Partial result data is reprinted from [67, p.7]

and $p_{iso} = 0$, dRL is incapable of resolving the crossing fibres of the numerical phantom for $\alpha \leq 60°$, with the resolvability problem becoming progressively worse with an increase in $p_{iso}$. In this case, the values of AAE effectively "mirror" those of $\alpha$ up to the point when dRL starts detecting the correct number of fibres, after which AAE becomes a decreasing function of the fibre crossing angle. It also deserves noting that for $b = 3000$ s/mm$^2$, dRL demonstrates considerably improved performance in terms of AAE, even surpassing CSD and Min-$\mathcal{L}_1$ for $\alpha \geq 50°$ and $p_{iso} \geq 0.5$. However, in all the alternative scenarios, the performance of dRL remains inferior to that of other methods under comparison. A possible explanation to this fact could be that the Lucy-Richardson algorithm exploited by dRL aims at recovering a maximum likelihood estimate under the assumption on measurement noise to be Poissonian. However, such a noise model can hardly be a good approximation to Rician distribution, which is inherent in MRI.

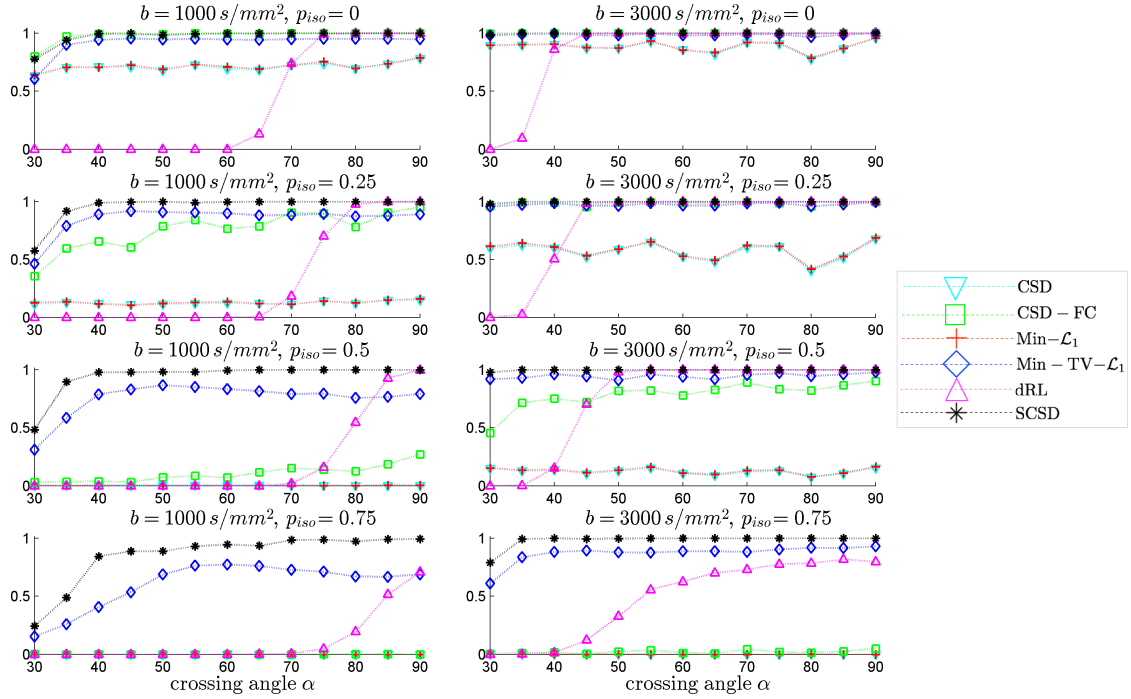


Figure 5.4: (Left columns of subplots) TP produced by the tested methods for different $\alpha$, $p_{iso}$, and $b = 1000$ s/mm$^2$; (Right columns of subplots) TP produced by the tested methods for different $\alpha$, $p_{iso}$, and $b = 3000$ s/mm$^2$. Partial result data is reprinted from [67, p.8]

Fig. 5.4 shows the values of TP which have been obtained using the proposed and reference methods for $b = 1000$ s/mm$^2$ (left column of subplots) and $b = 3000$ s/mm$^2$ (right column of subplots). One can see that, in the absence of isotropic diffusion (i.e., for $p_{iso} = 0$), the proposed SCSD method performs comparably to CSD-FC for both values of $b$ (with Min-TV-$\mathcal{L}_1$ being the next "best performer"). Yet, the moment $p_{iso}$ starts growing above zero, not only CSD and Min-$\mathcal{L}_1$, but also CSD-FC start loosing their ability to accurately estimate the true number of simulated fibres. Note that this is an expected result, considering the fact that neither of the above methods is endowed with facilities to explicitly account for the presence of isotropic diffusion, in which case over-estimation errors become inevitable (see below). At the same time, for $b = 3000$ s/mm$^2$, the proposed SCSD algorithm provides an ideal detection rate of one for $p_{iso} \in \{0, 0.25, 0.5\}$ and $\alpha \geq 30°$, as well as for $p_{iso} = 0.75$ and $\alpha \geq 35°$. It deserves noting that, owing to its properly accounting for the effect of isotropic diffusion, the performance of Min-TV-$\mathcal{L}_1$ is only marginally inferior to that of SCSD for $b = 3000$ s/mm$^2$ and $p_{iso} \leq 0.5$. One can also

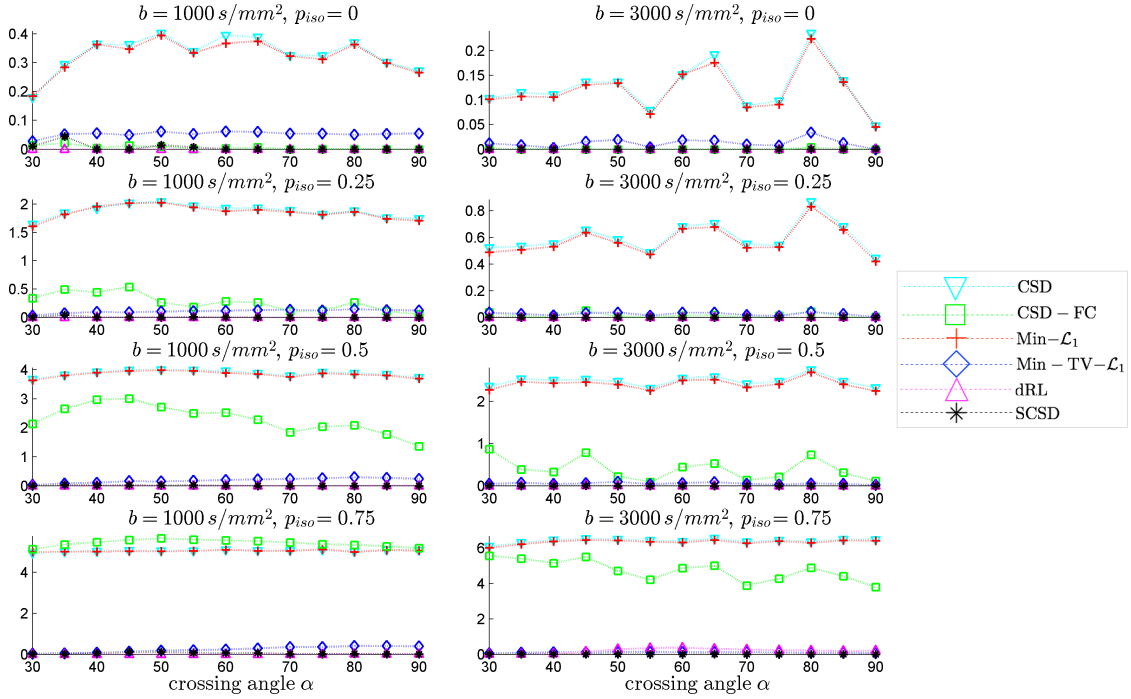


Figure 5.5: (Left columns of subplots) FP produced by the tested methods for different $\alpha$, $p_{iso}$, and $b = 1000$ s/mm$^2$; (Right columns of subplots) FP produced by the tested methods for different $\alpha$, $p_{iso}$, and $b = 3000$ s/mm$^2$.

see that a nearly ideal TP rate is reached by the dRL algorithm as well. However, it does not happen until after $\alpha$ reaches relatively large values (e.g., for $\alpha \geq 50^\circ$ with $p_{iso} = 0.5$ and $b = 3000$ s/mm$^2$).

One of the principle applications of HARDI is in multi-fibre tractography, in which case SD can be used to improve the resolvability of multiple fibre tracts within each given voxel in a region of interest (ROI). In this situation, overestimating the number of fibres is likely to produce spurious fibre tracts, thereby rendering the resulting reconstructions unreliable. Unfortunately, unless properly regularized, some SD routines tend to amplify the effect of noise, which in turn results in numerous false (local) maxima in reconstructed fODFs – the maxima that can be easily confused with the true modes of the latter. For this reason, it is important to compare the performance of SD routines in terms of the FP metric. Such comparative results are summarized in Fig. 5.5, the left and right columns of which correspond to the cases of $b = 1000$ s/mm$^2$ and $b = 3000$ s/mm$^2$, respectively. One

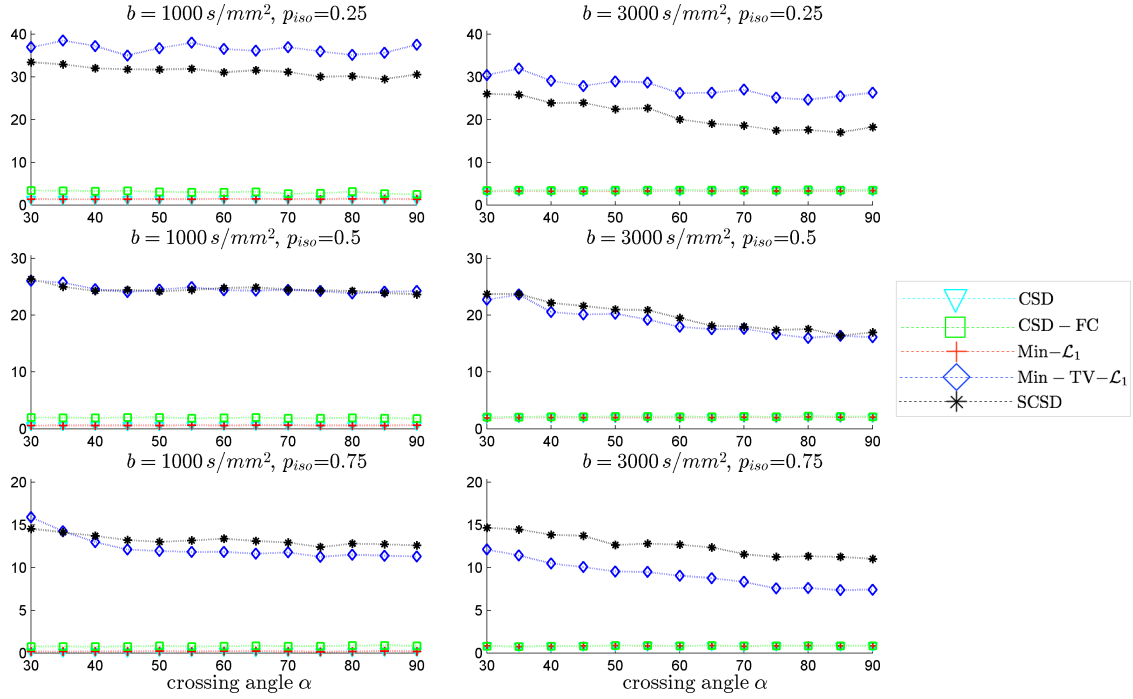


Figure 5.6: (Left columns of subplots) C produced by the tested methods for different $\alpha$, $p_{iso}$, and $b = 1000$ s/mm$^2$; (Right columns of subplots) C produced by the tested methods for different $\alpha$, $p_{iso}$, and $b = 3000$ s/mm$^2$.

can see that, in this case, the worst results are produced by CSD and Min-$\mathcal{L}_1$ for all values of $b$ and $\alpha$. The CSD-FC algorithm, on the other hand, provides a close to zero FP rate for $b = 3000$ s/mm$^2$ and $p_{iso} \in \{0, 0.25\}$. Unfortunately, its performance deteriorates for higher values of $p_{iso}$ (which is particularly noticeable for $b = 1000$ s/mm$^2$). Surprisingly, dRL seems to provide an ideal FP rate of zero for all values of $b$ and $\alpha$ under consideration. However, from our analysis of Fig. 5.4 it is not hard to see that it happens only because this method tends to underestimate the true number of simulated fibres. Finally, one can also see that, for $b = 3000$ s/mm$^2$, Min-TV-$\mathcal{L}_1$ provides a fairly small FP rate, whereas the proposed SCSD algorithm succeeds to attain the ideal FP rate of zero for all values of $p_{iso}$ and $\alpha$.

The results of our final quantitative comparison are summarized in Fig. 5.6, which shows the values of contrast $C$ obtained using different SD methods under comparison for $b = 1000$ s/mm$^2$ (left column of subplots) and $b = 3000$ s/mm$^2$ (right column of subplots).

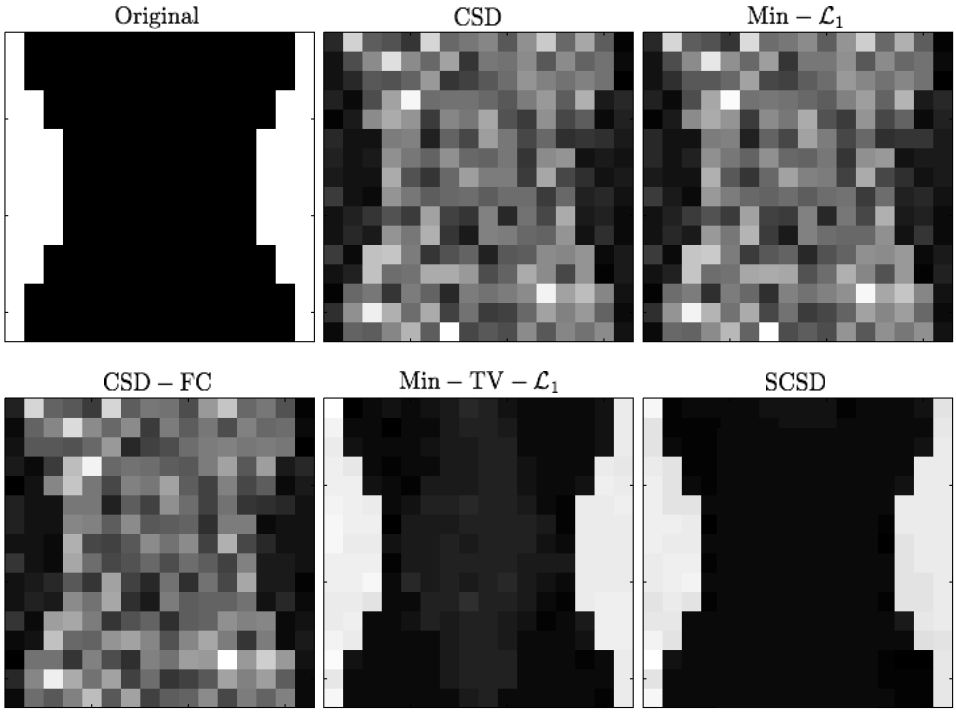


Figure 5.7: An "axial" slice of the IDMs recovered by different SD methods under comparison for $b = 3000$ s/mm$^2$, $\alpha = 45°$, and $p_{iso} = 0.5$. Partial result data is reprinted from [67, p.9]

Predictably enough, the best contrast is achieved by the Min-TV-$\mathcal{L}_1$ and SCSD algorithms, owing to their inherent ability to account for the presence of isotropic diffusion. Moreover, out of the two, the proposed SCSD algorithm yields the higher values of $C$ for all simulated scenarios. An additional illustration of the effect of incorporation and spatial regularization of the isotropic diffusion component is provided in Fig. 5.7, which depicts a 2-D "axial" slice of the IDMs reconstructed by different SD methods under comparison for $b = 3000$ s/mm$^2$, $\alpha = 45°$, and $p_{iso} = 0.5$. (Note that, for the sake of the clarity of visualization, the IDMs in Fig. 5.7 have been normalized so as to make their minimum and maximum values correspond to black and white pixel values, respectively.) One can see that the IDM reconstruction produced by SCSD is virtually indistinguishable from the original IDM (as shown in the upper, leftmost subplot of the figure), with the second best result produced by the Min-TV-$\mathcal{L}_1$ algorithm. At the same time, neither CSD, Min-$\mathcal{L}_1$ nor CSD-FC can attain a comparable accuracy of estimation of the spatial pattern of isotropic diffusion, as represented by the original IDM. In particular, even though their respective reconstructions

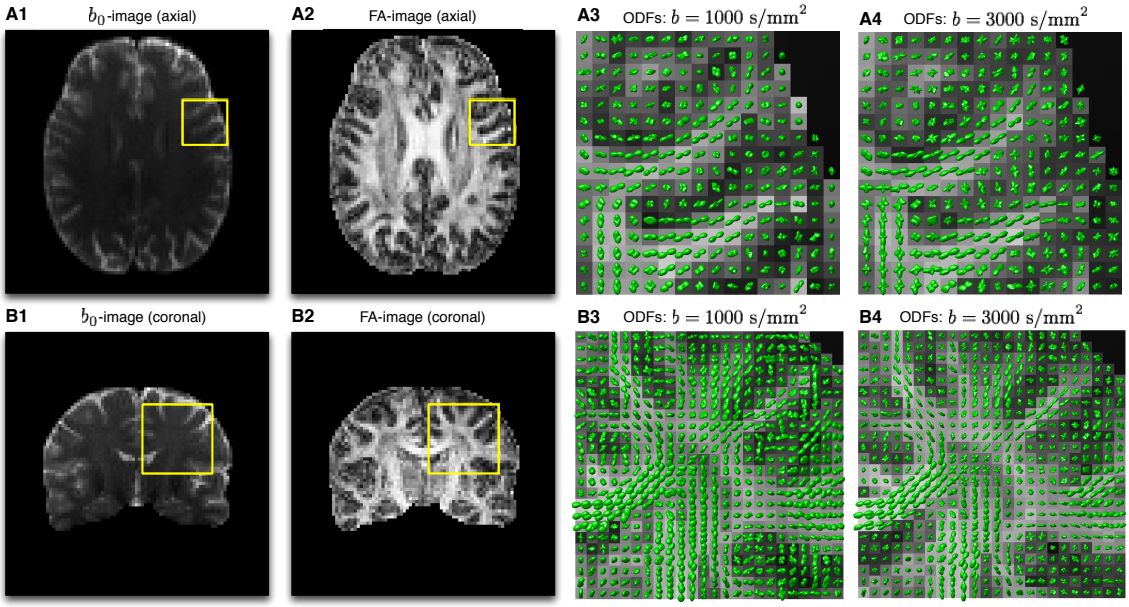


Figure 5.8: (Subplots A1-A2) $b_0$-image and FA images of an axial slice of an *in vivo* scan used for presentation of experimental results; (Subplots A3-A4) ODFs corresponding to the localized region indicated by the yellow squares in Subplots A1-A2 for $b = 1000$ s/mm$^2$ and $b = 3000$ s/mm$^2$, respectively; (Subplots B1-B2) $b_0$-image and FA images of a coronal slice from the same scan; (Subplots B3-B4) ODFs corresponding to the localized region indicated by the yellow squares in Subplots B1-B2 for $b = 1000$ s/mm$^2$ and $b = 3000$ s/mm$^2$, respectively. Reprinted from [67, p.9].

do bear some global resemblance to the original IDM, the level of estimation errors is too high to deem these reconstructions useful.

### 5.2.2 In Vivo Data Experiments

As the next step of our experimental study, real-life estimation has been performed using *in vivo* diffusion data (see Section 5.1.1). For the sake of the clarity of visualization, we restrict demonstration of the obtained reconstructions to 2-D views (aka "slices"), two examples of which are depicted in Fig 5.8. In particular, Subplots A1-A2 of the figure display an axial slice of the acquired $b_0$-volume and its corresponding fractional anisotropy (FA) image, respectively, while Subplots B1-B2 show a coronal slice from the same 3-D volume along with its associated FA image, in the same order. Additionally, Subplots

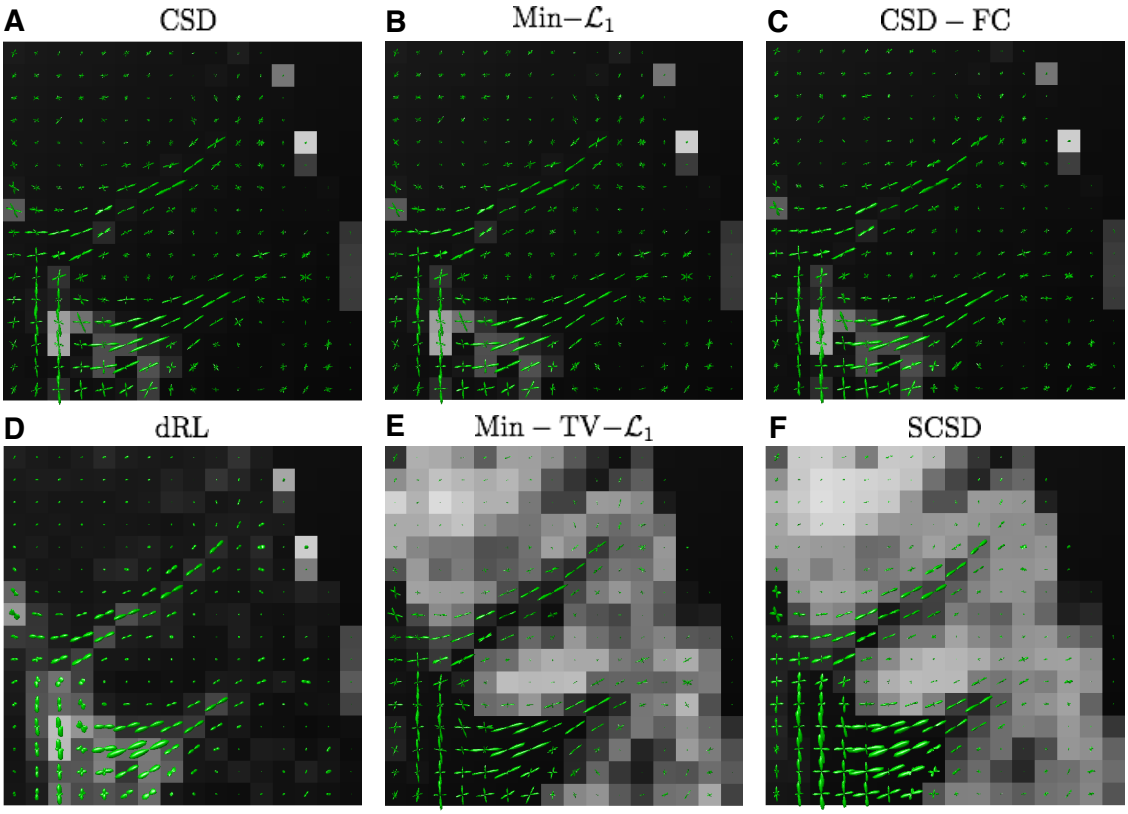


Figure 5.9: Estimated fODFs computed by CSD (A), Min-$\mathcal{L}_1$ (B), CSD-FC (C), dRL (D), Min-TV-$\mathcal{L}_1$ (E), and SCSD (F) for the axial view and $b = 1000$ s/mm$^2$. The fODFs are superimposed over the background of their related IDMs. Partial result data is reprinted from [67, p.10]

A3-A4 of Fig. 5.8 show the ODFs corresponding to the localized regions indicated by the yellow squares in Subplots A1-A2 for the case of $b = 1000$ s/mm$^2$ and $b = 3000$ s/mm$^2$, respectively. (To facilitate the analysis, the ODFs are shown over the background of their associated FA values.) Analogous results pertaining to the coronal view are shown in Subplots B3-B4 of the figure. It should be noted that, in both cases, the ODFs have been computed by means of the FRACT algorithm of [85], which seems to provide a reasonable balance between the robustness of more traditional QBI [27] and the high resolution gain of its solid-angle formulation [29].

Fig. 5.9 depicts the reconstructions of fODFs obtained using the CSD (Subplot A),

Min-$\mathcal{L}_1$ (Subplot B), CSD-FC (Subplot C), dRL (Subplot D), Min-TV-$\mathcal{L}_1$ (Subplot E), and SCSD (Subplot F) for the axial view and $b = 1000$ s/mm$^2$. As opposed to Subplots A3-4 and B3-4 in Fig. 5.8, the fODFs in Fig. 5.9 are shown superimposed over the values of their corresponding IDMs. Analysing these results reveals the principal drawback of SD methods which disregard the effect of isotropic diffusion. In particular, not only CSD and Min-$\mathcal{L}_1$ but also CSD-FC tend to yield spurious estimates of fODFs in anatomical regions corresponding to the cortical grey matter – the result which stands at odds with the fact that these regions are known to be devoid of neural fibre bundles. Further, although being capable of coping with the presence of isotropic diffusion, the dRL algorithm is "blending" $f_a$ and $f_{iso}$, which effectively impairs the angular resolution, and therefore the resolvability of crossing fibre tracts. Moreover, neither of the aforementioned methods has been found to be capable of reliably recovering the IDMs. At the same time, both Min-TV-$\mathcal{L}_1$ and SCSD yield anatomically consistent reconstructions of the IDMs, with much less noisy results obtained in the case of SCSD. (This point is further illustrated by Fig. 5.10 which shows the "zoomed-out" IDMs recovered by the SD methods under comparison for the case of $b = 1000$ s/mm$^2$.) Moreover, a closer inspection of Fig. 5.9 reveals that the fODFs estimated by means of SCSD are characterized by a smoother and more consistent spatial variability (owing to the fibre continuity constraint), thereby exhibiting a better adherence to the expected connectivity within an *in vivo* brain.

The reconstructions obtained for the same axial view and $b = 3000$ s/mm$^2$ are depicted in Fig. 5.11, whose composition is identical to that of Fig. 5.9. Although fairly close in appearance to the previous case, these reconstructions allow us to make a number of important observations. First of all, as compared to the case of $b = 1000$ s/mm$^2$, a wider bandwidth of the HARDI signals at $b = 3000$ s/mm$^2$ leads to a better angular resolution, which is particularly noticeable in the case of dRL. Moreover, despite considerably worse noise conditions, the fODF reconstructions obtained by means of Min-TV-$\mathcal{L}_1$ and SCSD have much less residual noise over the areas occupied by cortical grey matter, where isotropic diffusion is expected to prevail. This fact indicates the effectiveness of the regularization schemes exploited by these SD methods. Finally, a closer inspection of the glyphs in Fig. 5.11 reveals that the fODF reconstructions yielded by SCSD demonstrate a better spatial smoothness and anatomical consistency. The same observations can be made in the case of the coronal view, as shown in Figs. 5.12 and 5.13 for $b = 1000$ s/mm$^2$ and $b = 3000$ s/mm$^2$, respectively. (Note that the local region represented by these figures has been chosen according to the results in [42], which also provides indication of specific fibre bundles within the selected regions of interest).

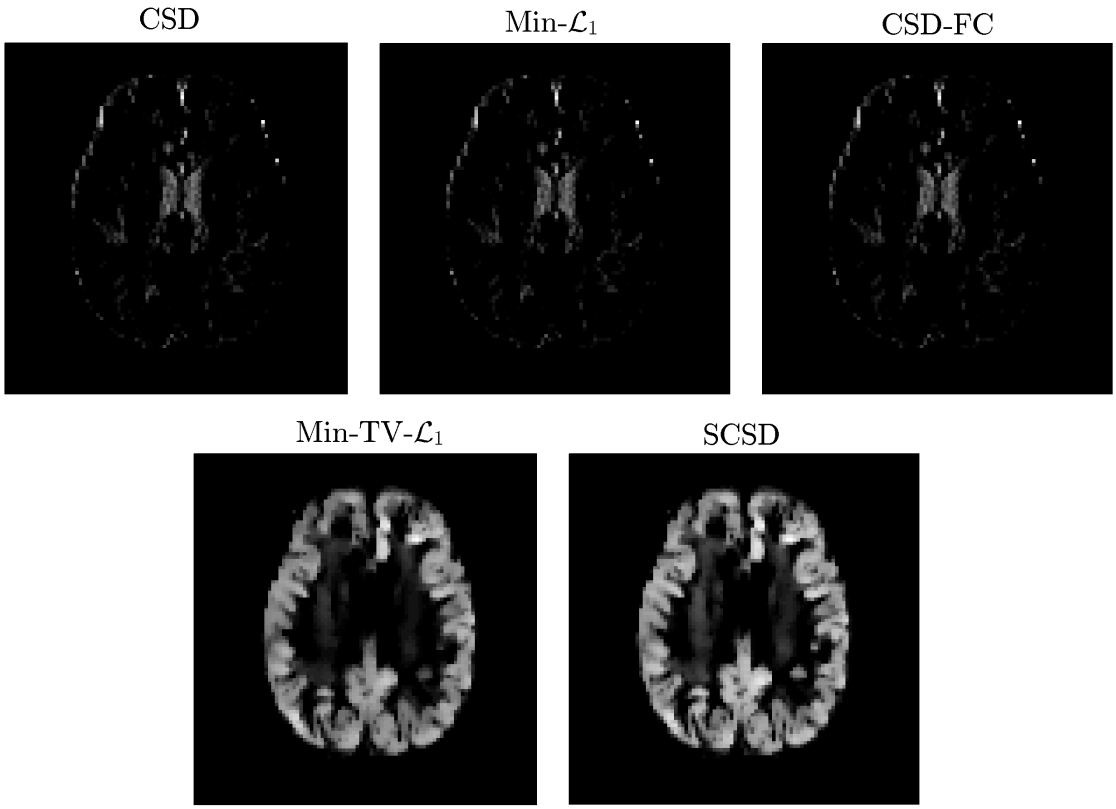


Figure 5.10: "Axial" IDMs estimated by various SD methods for $b = 1000$ s/mm$^2$. Partial result data is reprinted from [67, p.11]

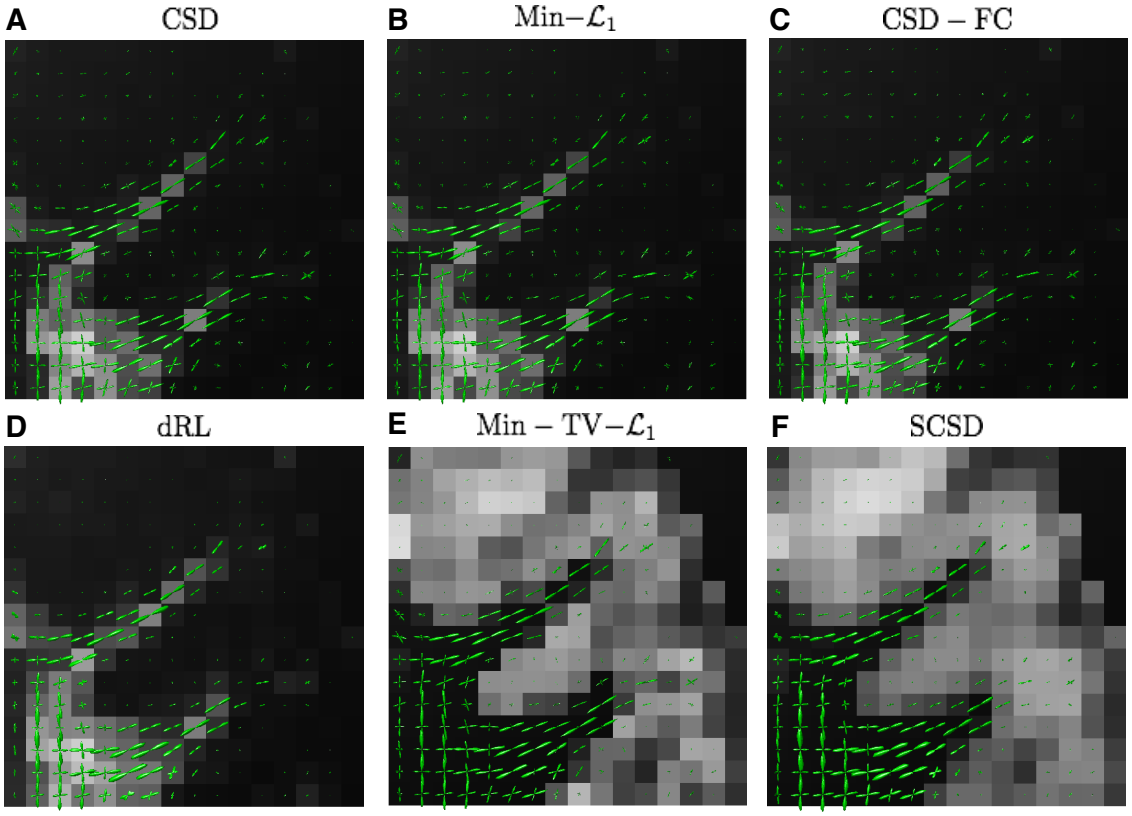


Figure 5.11: Estimated fODFs computed by CSD (A), Min-$\mathcal{L}_1$ (B), CSD-FC (C), dRL (D), Min-TV-$\mathcal{L}_1$ (E), and SCSD (F) for the axial view and $b = 3000$ s/mm$^2$. The fODFs are superimposed over the background of their related IDMs. Partial result data is reprinted from [67, p.10]

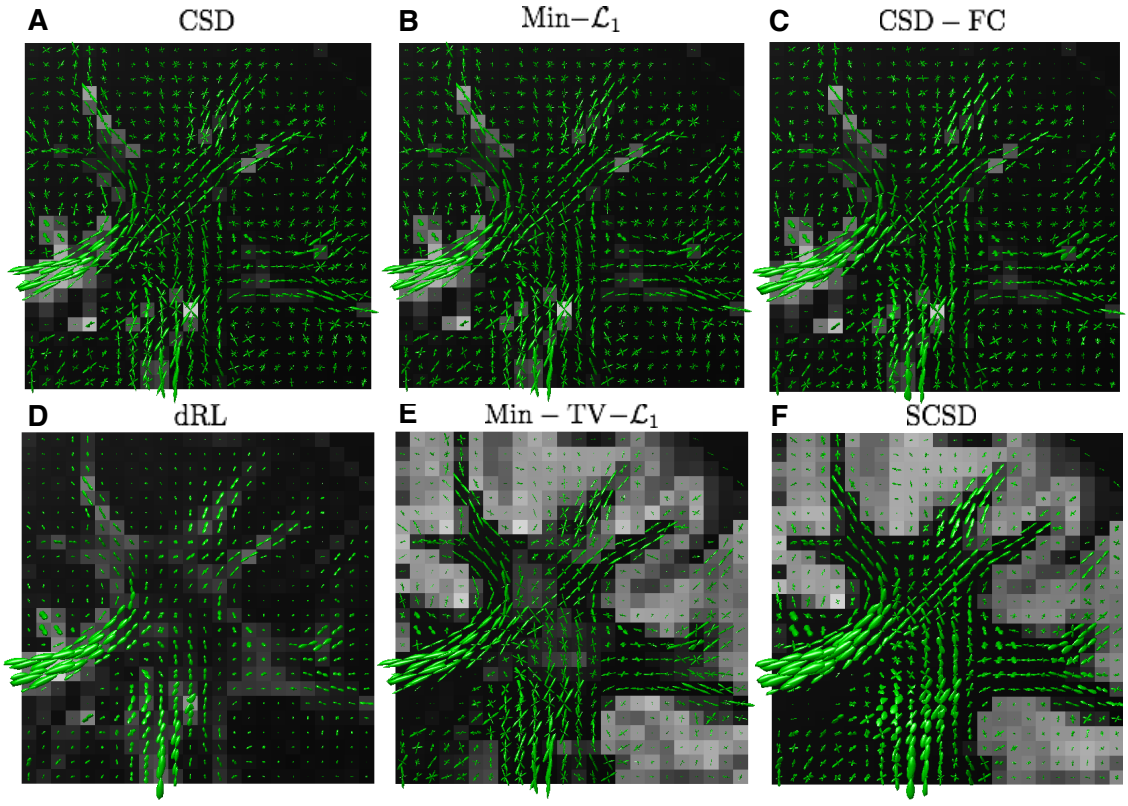


Figure 5.12: Estimated fODFs computed by CSD (A), Min-$\mathcal{L}_1$ (B), CSD-FC (C), dRL (D), Min-TV-$\mathcal{L}_1$ (E), and SCSD (F) for the coronal view and $b = 1000$ s/mm$^2$. The fODFs are superimposed over the background of their related IDMs. Partial result data is reprinted from [67, p.11]

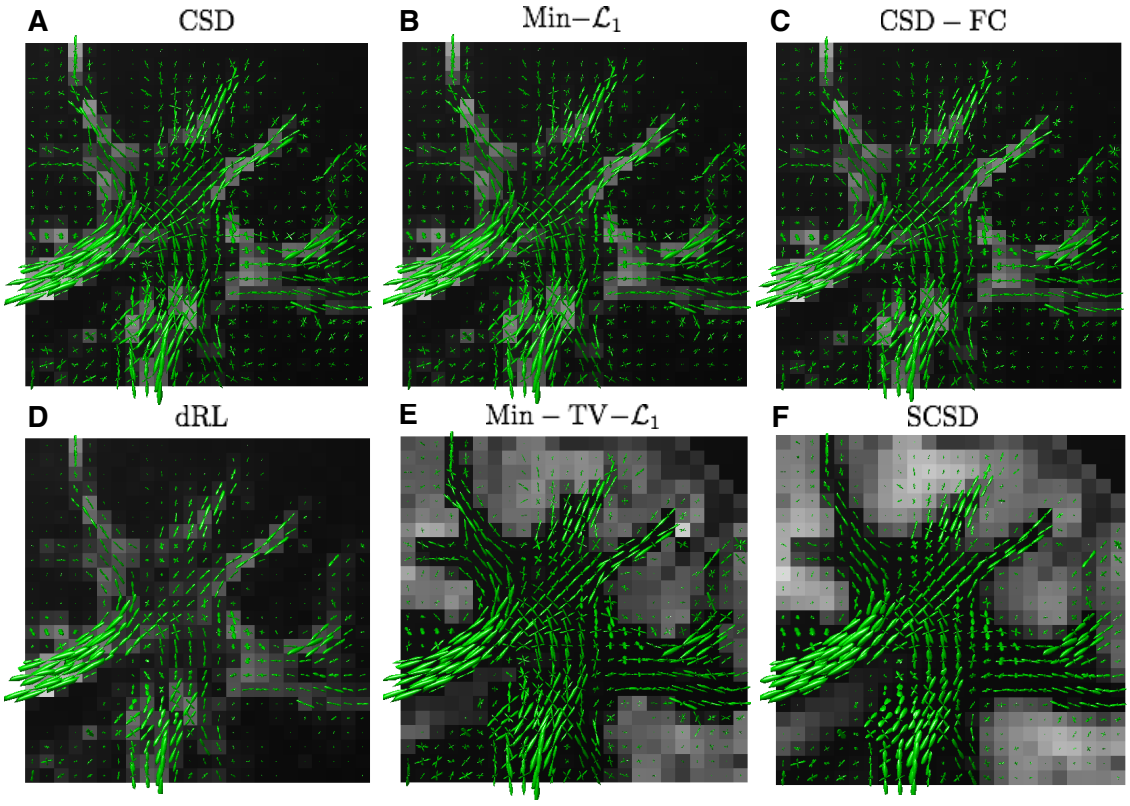


Figure 5.13: Estimated fODFs computed by CSD (A), Min-$\mathcal{L}_1$ (B), CSD-FC (C), dRL (D), Min-TV-$\mathcal{L}_1$ (E), and SCSD (F) for the coronal view and $b = 3000$ s/mm$^2$. The fODFs are superimposed over the background of their related IDMs. Partial result data is reprinted from [67, p.11]

# Chapter 6

# Summary and Conclusion

In the present thesis, we introduce a novel approach to the problem of non-blind SD of HARDI signals. As a derivation of early proposed SD method Min-TV-$\mathcal{L}_1$ [67], the new SD method preserves the advantages of Min-TV-$\mathcal{L}_1$, which are reiterated here. As opposed to many alternative methods of SD, the proposed algorithm can explicitly account for the effect of isotropic diffusion, which makes it capable of performing reliably across the entire brain, thereby avoiding the need to restrict the computations to the regions of white matter. In particular, in addition to reconstruction of fODFs $f_a(\mathbf{u}|\mathbf{r})$, the algorithm can also yield a useful estimation of its related IDM $f_{iso}(\mathbf{r})$, which quantifies a relative contribution of the isotropic diffusion component as well as its spatial pattern. Moreover, one of the principal contributions of the present work is to demonstrate the effectiveness of exploiting different prior models for regularization of the spatial-domain behaviour of the reconstructed fODFs and IDMs. Specifically, the fibre continuity model of [50] has been used to force the local maxima of the fODFs to vary consistently throughout the brain, whereas the bounded variation model of [80] has helped us to achieve piecewise smooth reconstructions of the IDMs (which appear to be in a good agreement with brain anatomy).

The estimated IDMs could serve as an additional input to fibre tractography or an independent observation for further statistical inference. In both scenarios, we believe the spatial regularization is essential to make the reconstructions of fODFs and IDMs be in a better agreement with the anatomical structure of the brain. Furthermore, as long as fibre tractography is concerned, our experimental results indicate the possibility to achieve a better angular resolution using higher values of $b$ (e.g., $b = 3000$ s/mm$^2$). It is important to note that this outcome would not have been possible without regularizing the IDMs (as explained in Section 4.3 and experimentally verified in Section 5.2) due to the adverse effect of measurement noises.

The proposed SCSD algorithm has been formulated as a convex minimization problem, which admits a unique and stable minimizer. Moreover, an important contribution of this work is to introduce a computationally efficient scheme for numerical implementation of SCSD. In particular, using ADMM, we have been able to find the optimal solution via a sequence of simpler optimization problems, which are both computationally efficient and amenable to parallel computations.

# Chapter 7

# Future Work

In the proposed SCSD method, two spatial regularizers play a vital role in constraining the reconstructed fODF and IDM. Specifically, the fibre continuity model of [50] has been used to force the local maxima of the fODFs to vary consistently throughout the brain, whereas the bounded variation model of [80] has helped us to achieve piecewise smooth reconstructions of the IDMs (which appear to be in a good agreement with brain anatomy). Moreover, although the results of our experimental study support the above prior models, the latter are by no means exclusive, and hence additional improvements could still be achieved through the use of more advanced methods of spatial regularization. Thus, for example, the bounded variation model could have been replaced by requiring the diffusion-encoded images $s_k(\mathbf{r})$ to have sparse representation coefficients in the domain of a multi-resolution transform (e.g., framelets [86]). At the same time, the linear filtering resulting from the fibre continuity assumption (as explained in the Appendix) could have been potentially replaced by a non-linear spatially-adaptive filtering. Whether or not the above modifications can result in substantial improvements in the quality of SD reconstruction, it remains a subject of our further research.

The performance of the SCSD method has been quantitatively compared against that of a number of alternative approaches, which have been specially selected to demonstrate the importance of various assumptions and constraints. Although SCSD has been shown to outperform the reference methods in terms of all the comparison metrics used, it still has a few limitations, whose mitigation could potentially produce more accurate reconstructions. Thus, from a Bayesian perspective, the cost functional in (4.15) suggests the measurement noise to be of an additive Gaussian nature, while in practice it is more likely to be Rician. It should be noted, however, that the Gaussian model can serve as a good approximation to the Rician model for relative high values of SNR (i.e., SNR $\gg 5$) [87]. Besides,

using the Gaussian model is essential for rendering the optimization problem convex and computationally tractable. Yet, we note that, if using the Rician statistics is preferred, it should be straightforward to modify the SCSD algorithm as detailed in [87] for a similar setting. Finally, it should also be mentioned that the present version of SCSD assumes the diffusion outside white matter to be purely isotropic, while evidence exists that the diffusion within grey matter could exhibit a certain degree of anisotropy [88]. To account for this phenomenon, one could, for instance, replace the constant column of $\Phi$ in (4.7) by spherical harmonics up to second degree inclusive. Exploring such a modification and its effect on the quality of SD reconstruction defines another direction of our future research.

Another direction of SCSD improvement is to make it adapt to multishell signals in **q**-space [13, 14, 15], illustrated in Fig. 2.4 (d). As discussed in section 3.2.4, the diffusion signals sampled on **q**-space shell with high **q** value can bring the benefit of higher angular resolution but along with the contamination of high noise component, meanwhile low **q** value shell possesses high SNR but lower angular resolution. Combining different **q**-space shells together and properly selecting model and regularizers can make full use of their advantages and avoid the shortcomings.

# APPENDICES

# Appendix A

# Directional Low-Pass Filter

To find a computationally efficient way to solve (4.26), it is convenient to slightly simplify the notations first. To this end, we note that the problem has a general form of

$$\min_{w} \left\{ \frac{1}{2} \|w - q\|_2^2 + \tau \|\mathbf{T}_{\mathbf{v}} w\|_2^2 \right\}, \tag{A.1}$$

with $w$ interpreted as an $I$-dimensional (row) vector to be optimized over, $q$ is a data vector, $\tau > 0$ is a fixed regularization constant, and $\mathbf{T}_{\mathbf{v}} : \mathbb{R}^I \to \mathbb{R}^I$ is the operator of directional differencing in the direction of $\mathbf{v} \in \mathbb{S}^2$. The problem (A.1) is a standard LS problem, whose optimal solution $w^\star$ satisfies a system of normal equations of the form

$$\left(\mathbf{I} + 2\tau\, \mathbf{T}_{\mathbf{v}}^* \mathbf{T}_{\mathbf{v}}\right) w^\star = q, \tag{A.2}$$

with $\mathbf{T}_{\mathbf{v}}^*$ standing for the adjoint of $\mathbf{T}_{\mathbf{v}}$.

It goes without saying that a straightforward inversion of the matrix on the left-hand side of (A.2) cannot be accepted as a practical option, even for moderately sized diffusion-encoded images. Thus, a more efficient way to solve (A.2) needs to be found. To this end, we first note that, if the practical computation of partial differences was based on a standard backward-differencing scheme, then multiplication by $\mathbf{T}_{\mathbf{v}}$ in the spatial domain would be equivalent to linear filtering with frequency response $H_{\mathbf{v}}(\boldsymbol{\omega})$ given by

$$H_{\mathbf{v}}(\boldsymbol{\omega}) = 2 \sum_{d=1}^{3} \mathbf{v}_d \sin \frac{\boldsymbol{\omega}_d}{2} e^{-\jmath\left(\frac{\boldsymbol{\omega}_d - \pi}{2}\right)}, \quad \boldsymbol{\omega} \in \mathbb{R}^3, \tag{A.3}$$

with $\mathbf{v} = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ and $\boldsymbol{\omega} = (\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \boldsymbol{\omega}_3)$. In this case, by letting $\hat{w}^\star(\boldsymbol{\omega})$ and $\hat{q}(\boldsymbol{\omega})$ denote the discrete-space Fourier transform (DSFT) of $w$ and $q$, respectively, the optimal solution in (A.2) could be defined in the DSFT domain as

$$\hat{w}^\star(\boldsymbol{\omega}) = \left(1 + 2\tau |H_{\mathbf{v}}(\boldsymbol{\omega})|^2\right)^{-1} \hat{q}(\boldsymbol{\omega}), \tag{A.4}$$

which suggests that $w^\star(\boldsymbol{\omega})$ is, in fact, a linearly filtered version of $q$.

Although it is definitely possible to use (A.4) in practical computations, such an approach would not be recommended for two main reasons. First, FFT-based implementation of linear filtering would entail the use of periodic boundary conditions, which might not be natural for the case at hand. Second, the logarithmic complexity of FFT might still be considered to be prohibitively high (especially taking into account the fact that the filter needs to be applied $J$ times for $J$ different values of $\mathbf{v}$).

A practical alternative to an FFT-based computation of $\hat{w}^\star(\boldsymbol{\omega})$ could be first to transform the frequency response $(1 + 2\tau |H_{\mathbf{v}}(\boldsymbol{\omega})|^2)^{-1}$ back into the spatial domain, followed by truncating the impulse response thus obtained. In our experiments, the impulse responses have been truncated to a size of $7 \times 7 \times 7$ voxels, with an associated error being less than 1% for a typical choice of the parameter $\tau$ (or, equivalently, $\mu$). The resulting finite impulse response filters have distinctive (directional) low-pass characteristics, and they can be easily applied with a linear complexity under arbitrary boundary conditions.

# References


[1] W. Hendee and E. Ritenour, *Medical Imaging Physics*, 4th ed. New York, NY, USA: John Wiley & Sons, Inc., Oct. 2003.

[2] T. E. E. Johansen-Berg, H., & Behrens, *Diffusion MRI: From Quantitative Measurement to In-Vivo Neuroanatomy*, 1st ed. Salt Lake City, UT, USA: Academic Press, 2009.

[3] P. J. Basser, J. Mattiello, and D. LeBihan, "Estimation of the Effective Self-Diffusion Tensor from the NMR Spin Echo." *Journal of Magnetic Resonance. Series B*, vol. 103, no. 3, pp. 247–254, Mar. 1994.

[4] ——, "MR Diffusion Tensor Spectroscopy and Imaging." *Biophysical Journal*, vol. 66, no. 1, pp. 259–267, Jan. 1994.

[5] D. L. Bihan, J.-F. Mangin, C. Poupon, C. Clark, S. Pappata, N. Molko, and H. Chabriat, "Diffusion Tensor Imaging: Concepts and Applications," *Journal of Magnetic Resonance Imaging*, vol. 13, pp. 534–546, 2001.

[6] D. Le Bihan, E. Breton, D. Lallemand, P. Grenier, E. Cabanis, and M. Laval-Jeantet, "MR Imaging of Intravoxel Incoherent Motions: Application to Diffusion and Perfusion in Neurologic Disorders." *Radiology*, vol. 161, no. 2, pp. 401–407, Nov. 1986.

[7] S. Mori, *Introduction to Diffusion Tensor Imaging*, 1st ed. Oxford, UK: Elsevier, 2007.

[8] L. Frank, "Anisotropy in High Angular Resolution Diffusion-Weighted MRI," *Magnetic Resonance in Medicine*, vol. 45, pp. 935–939, 2001.

[9] L. R. Frank, "Characterization of Anisotropy in High Angular Resolution Diffusion-Weighted MRI," *Magnetic Resonance in Medicine*, vol. 47, no. 6, pp. 1083–1099, 2002.

[10] D. S. Tuch, T. G. Reese, M. R. Wiegell, N. Makris, J. W. Belliveau, and V. J. Wedeen, "High Angular Resolution Diffusion Imaging Reveals Intravoxel White Matter Fiber Heterogeneity." *Magnetic Resonance in Medicine*, vol. 48, no. 4, pp. 577–582, Oct. 2002.

[11] D. Tuch, "Diffusion MRI of Complex Tissue Structure," Ph.D. dissertation, MIT Division of Health Sciences and Technology, Harvard University, 2002.

[12] V. J. Wedeen, P. Hagmann, W.-Y. I. Tseng, T. G. Reese, and R. M. Weisskoff, "Mapping Complex Tissue Architecture with Diffusion Spectrum Magnetic Resonance Imaging." *Magnetic Resonance in Medicine*, vol. 54, no. 6, pp. 1377–1386, Dec. 2005.

[13] M. Descoteaux, R. Deriche, D. Le Bihan, J.-F. Mangin, and C. Poupon, "Multiple Q-Shell Diffusion Propagator Imaging." *Medical Image Analysis*, vol. 15, no. 4, pp. 603–621, Aug. 2011.

[14] M. I. Menzel, E. T. Tan, K. Khare, J. I. Sperl, K. F. King, X. Tao, C. J. Hardy, and L. Marinelli, "Accelerated Diffusion Spectrum Imaging in the Human Brain Using Compressed Sensing," *Magnetic Resonance in Medicine*, vol. 66, pp. 1226–1233, 2011.

[15] S. Merlet and D. Rachid, "Compressed Sensing for Accelerated EAP Recovery in Diffusion MRI," in *Medical Image Computing and Computer-Assisted Intervention*, Beijing, China, 2010, pp. 14–21.

[16] Y. Assaf, R. Z. Freidlin, G. K. Rohde, and P. J. Basser, "New Modeling and Experimental Framework to Characterize Hindered and Restricted Water Diffusion in Brain White Matter." *Magnetic Resonance in Medicine*, vol. 52, no. 5, pp. 965–978, Nov. 2004.

[17] S. Merlet, J. Cheng, A. Ghosh, and R. Deriche, "Spherical Polar Fourier EAP and ODF Reconstruction via Compressed Sensing in Diffusion MRI," in *Biomedical Imaging: From Nano to Macro*, Chicago, IL, USA, 2011, pp. 365–371.

[18] Y.-C. Wu and A. L. Alexander, "Hybrid Diffusion Imaging," *NeuroImage*, vol. 36, pp. 617–629, 2007.

[19] J. H. Jensen, J. A. Helpern, A. Ramani, H. Lu, and K. Kaczynski, "Diffusional Kurtosis Imaging: The Quantification of Non-Gaussian Water Diffusion by Means of Magnetic Resonance Imaging," *Magnetic Resonance in Medicine*, vol. 53, pp. 1432–1440, 2005.

[20] H. Zhang, T. Schneider, C. A. Wheeler-Kingshott, and D. C. Alexander, "NODDI: Practical In-Vivo Neurite Orientation Dispersion and Density Imaging of the Human Brain," *NeuroImage*, vol. 61, no. 4, pp. 1000–1016, 2012.

[21] H. Assemlal, D. Tschumperle, and L. Brun, "Efficient Computation of PDF-Based Characteristics from Diffusion MR Signal," in *Medical Image Computing and Computer-Assisted Intervention*, New York, USA, 2008, pp. 70–78.

[22] E. Ozarslan, C. G. Koay, T. M. Shepherd, M. E. Komlosh, M. O. Irfanoglu, and C. P. P. J. Basser, "Mean Apparent Propagator (MAP) MRI: A Novel Diffusion Imaging Method for Mapping Tissue Microstructure," *NeuroImage*, vol. 78, pp. 16–32, 2013.

[23] D. S. Tuch, T. G. Reese, M. R. Wiegell, and V. J. Wedeen, "Diffusion MRI of Complex Neural Architecture." *Neuron*, vol. 40, no. 5, pp. 885–895, Dec. 2003.

[24] D. S. Tuch, "Q-Ball Imaging." *Magnetic Resonance in Medicine*, vol. 52, no. 6, pp. 1358–1372, Dec. 2004.

[25] D. C. Alexander, "Multiple-Fiber Reconstruction Algorithms for Diffusion MRI," *Annals of the New York Academy of Science*, vol. 1064, pp. 113–133, 2005.

[26] C. P. Hess, P. Mukherjee, E. T. Han, D. Xu, and D. R. Vigneron, "Q-Ball Reconstruction of Multimodal Fiber Orientations Using the Spherical Harmonic Basis," *Magnetic Resonance in Medicine*, vol. 56, pp. 104–117, 2006.

[27] M. Descoteaux, E. Angelino, S. Fitzgibbons, and R. Deriche, "Regularized, Fast, and Robust Analytical Q-Ball Imaging," *Magnetic Resonance in Medicine*, vol. 58, pp. 497–510, 2007.

[28] ——, "Apparent Diffusion Coefficients from High Angular Resolution Diffusion Images: Estimation and Applications," *Magnetic Resonance in Medicine*, vol. 56, no. 2, pp. 395–410, 2006.

[29] I. Aganj, C. Lenglet, G. Sapiro, E. Yacoub, K. Ugurbil, and N. Harel, "Reconstruction of the Orientation Distribution Function in Single- and Multiple-Shell Q-Ball Imaging within Constant Solid Angle," *Magnetic Resonance in Medicine*, vol. 64, no. 2, pp. 554 – 566, 2010.

[30] A. Barnett, "Theory of Q-Ball Imaging Redux: Implications for Fiber Tracking," *Magnetic Resonance in Medicine*, vol. 62, pp. 910–923, 2009.

[31] J. Malcolm, M. Shenton, and Y. Rathi, "Neural Tractography Using an Unscented Kalman Filter," in *International Precious Metals Institute*, Williamsburg, VA, USA, 2009, pp. 126–138.

[32] D. M. H. Jr., H. Hendriks, and P. T. Kim, "Spherical Deconvolution," *Journal of Multivariate Analysis*, vol. 67, pp. 1–22, 1998.

[33] J.-D. Tournier, F. Calamante, D. G. Gadian, and A. Connelly, "Direct Estimation of the Fiber Orientation Density Function from Diffusion-Weighted MRI Data Using Spherical Deconvolution." *NeuroImage*, vol. 23, no. 3, pp. 1176–1185, Nov. 2004.

[34] J.-D. Tournier, F. Calamante, and A. Connelly, "Robust Determination of the Fibre Orientation Distribution in Diffusion MRI: Non-Negativity Constrained Super-Resolved Spherical Deconvolution." *NeuroImage*, vol. 35, no. 4, pp. 1459–1472, May 2007.

[35] J.-D. Tournier, C.-H. Yeh, F. Calamante, K.-H. Cho, A. Connelly, and C.-P. Lin, "Resolving Crossing Fibres Using Constrained Spherical Deconvolution: Validation Using Diffusion-Weighted Imaging Phantom Data." *NeuroImage*, vol. 42, no. 2, pp. 617–625, Aug. 2008.

[36] A. W. Anderson, "Measurement of Fiber Orientation Distributions Using High Angular Resolution Diffusion Imaging," *Magnetic Resonance in Medicine*, vol. 54, pp. 1194–1206, 2005.

[37] K. E. Sakaie and M. J. Lowe, "An Objective Method for Regularization of Fiber Orientation Distributions Derived from Diffusion-Weighted MRI." *NeuroImage*, vol. 34, no. 1, pp. 169–176, Jan. 2007.

[38] D. C. Alexander, "Maximum Entropy Spherical Deconvolution for Diffusion MRI." in *Information Processing in Medical Imaging*, vol. 19, Glenwood Springs, CO, USA, Jan. 2005, pp. 76–87.

[39] M. Jing, T. M. McGinnity, S. Coleman, H. Zhang, A. Fuchs, and J. A. S. Kelso, "Enhancement of Fiber Orientation Distribution Reconstruction in Diffusion-Weighted Imaging by Single Channel Blind Source Separation," *IEEE Transactions on Medical Imaging*, vol. 59, no. 2, pp. 363–373, 2012.

[40] B. Jian and B. C. Vemuri, "A Unified Computational Framework for Deconvolution to Reconstruct Multiple Fibers from Diffusion Weighted MRI." *IEEE Transactions on Medical Imaging*, vol. 26, no. 11, pp. 1464–1471, Nov. 2007.

[41] F. Dell'Acqua, G. Rizzo, P. Scifo, R. A. Clarke, G. Scotti, and F. Fazio, "A Model-Based Deconvolution Approach to Solve Fiber Crossing in Diffusion-Weighted MR Imaging." *IEEE Transactions on Bio-Medical Engineering*, vol. 54, no. 3, pp. 462–472, Mar. 2007.

[42] F. Dell'acqua, P. Scifo, G. Rizzo, M. Catani, A. Simmons, G. Scotti, and F. Fazio, "A Modified Damped Richardson-Lucy Algorithm to Reduce Isotropic Background Effects in Spherical Deconvolution." *NeuroImage*, vol. 49, no. 2, pp. 1446–1458, Jan. 2010.

[43] T. Schultz and H.-P. Seidel, "Estimating Crossing Fibers: A Tensor Decomposition Approach." *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 6, pp. 1635–1642, 2008.

[44] P. Yap and D. Shen, "Spatial Transformation of DWI Data Using Non-Negative Sparse Representation." *IEEE Transactions on Medical Imaging*, vol. 31, no. 11, pp. 2035–2049, Jun. 2012.

[45] E. Kaden, A. Anwander, and T. R. Knösche, "Variational Inference of the Fiber Orientation Density Using Diffusion MR Imaging." *NeuroImage*, vol. 42, no. 4, pp. 1366–1380, Oct. 2008.

[46] E. Kaden and F. Kruggel, "Nonparametric Bayesian Inference of the Fiber Orientation Distribution from Diffusion-Weighted MR Images." *Medical Image Analysis*, vol. 16, no. 4, pp. 876–888, May 2012.

[47] A. Goh, C. Lenglet, P. M. Thompson, and R. Vidal, "Estimating Orientation Distribution Functions with Probability Density Constraints and Spatial Regularity." in *Medical Image Computing and Computer-Assisted Intervention*, vol. 12, no. Pt 1, London, UK, Jan. 2009, pp. 877–885.

[48] A. Raj, C. Hess, and P. Mukherjee, "Spatial HARDI: Improved Visualization of Complex White Matter Architecture with Bayesian Spatial Regularization," *NeuroImage*, vol. 54, pp. 396–409, 2011.

[49] A. Ramirez-Manzanares, M. Rivera, B. C. Vemuri, P. Carney, and T. Mareci, "Diffusion Basis Functions Decomposition for Estimating White Matter Intravoxel

Fiber Geometry." *IEEE Transactions on Medical Imaging*, vol. 26, no. 8, pp. 1091–1102, Aug. 2007.

[50] M. Reisert and V. G. Kiselev, "Fiber Continuity: An Anisotropic Prior for ODF Estimation." *IEEE Transactions on Medical Imaging*, vol. 30, no. 6, pp. 1274–1283, Jun. 2011.

[51] P. L. Combettes and J.-C. Pesquet, "Proximal Splitting Methods in Signal Processing," in *Springer Optimization and Its Applications*. New York NY, USA: Springer, 2011, vol. 49, pp. 185–212.

[52] T. Moritani, S. Ekholm, and P.-L. Westesson, *Diffusion-Weighted MR Imaging of the Brain*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005.

[53] D.-M. Koh and H. C. Thoeny, *Diffusion-Weighted MR Imaging: Applications in the Body*. London, UK: Springer-Verlag Berlin Heidelberg, 2010.

[54] D. Weishaupt, V. D. Köchli, and B. Marincek, *How Does MRI Work?: An Introduction to the Physics and Function of Magnetic Resonance Imaging*. New York, NY, USA: Springer, 2006.

[55] L. Schroder and C. Faber, *In Vivo NMR Imaging: Methods and Protocols*. London, UK: Springer New York Dordrecht Heidelberg London, 2011.

[56] K. K. Shung, M. B. Smith, and B. M. W. Tsui, *Principles of Medical Imaging*. Salt Lake City, UT, USA: Academic Press, 1992.

[57] D. W. McRobbie, E. A. Moore, M. J. Graves, and M. R. Prince, *MRI from Picture to Proton*, 2nd ed. Cambridge, UK: Cambridge University Press, 2007.

[58] E. O. Stejskal and J. E. Tanner, "Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient," *The Journal of Chemical Physics*, vol. 42, no. 1, p. 288, 1965.

[59] E. O. Stejskal, "Use of Spin Echoes in a Pulsed Magnetic-Field Gradient to Study Anisotropic, Restricted Diffusion and Flow," *The Journal of Chemical Physics*, vol. 43, no. 10, p. 3597, May 1965.

[60] Wedeen VJ, R. TG, T. DS, M. Weigell, J.-G. Dou, W. RM, and D. Chesler, "Mapping Fiber Orientation Spectra in Cerebral White Matter with Fourier Transform Diffusion MRI," in *International Society for Magnetic Resonance in Medicine*, Denver, CO, USA, 2000, p. 82.

[61] D. Tuch, R. Weisskoff, J. Belliveau, and V. Wedeen, "High Angular Resolution Diffusion Imaging of the Human Brain," in *International Society for Magnetic Resonance in Medicine*, Philadelphia, PA, USA, 1999.

[62] T. E. J. Behrens, H. J. Berg, S. Jbabdi, M. F. S. Rushworth, and M. W. Woolrich, "Probabilistic Diffusion Tractography with Multiple Fibre Orientations: What Can We Gain?" *NeuroImage*, vol. 34, no. 1, pp. 144–155, Jan. 2007.

[63] T. Hosey, G. Williams, and R. Ansorge, "Inference of Multiple Fiber Orientations in High Angular Resolution Diffusion Imaging." *Magnetic Resonance in Medicine*, vol. 54, no. 6, pp. 1480–1489, Dec. 2005.

[64] A. Tristan-Vega, C. F. Westin, and S. Aja-Fernandez, "Estimation of Fiber Orientation Probability Density Functions in High Angular Resolution Diffusion Imaging," *NeuroImage*, vol. 47, pp. 638–650, 2009.

[65] B. Jian and B. C. Vemuri, "Multi-Fiber Reconstruction from Diffusion MRI Using Mixture of Wisharts and Sparse Deconvolution." *Information Processing in Medical Imaging*, vol. 20, pp. 384–395, Jan. 2007.

[66] F.-C. Yeh, V. J. Wedeen, and W.-Y. I. Tseng, "Estimation of Fiber Orientation and Spin Density Distribution by Diffusion Deconvolution." *NeuroImage*, vol. 55, no. 3, pp. 1054–1062, Apr. 2011.

[67] Q. Zhou, O. Michailovich, and Y. Rathi, "Resolving Complex Fibre Architecture by Means of Sparse Spherical Deconvolution in the Presence of Isotropic Diffusion," in *SPIE Medical Imaging*, San Diego, CA, USA, 2014.

[68] P. Funk, "Über Eine Geometrische Anwendung Der Abelschen Integralgleichung," *Mathematische Annalen*, vol. 77, no. 1, pp. 129–135, Mar. 1915.

[69] D. L. Bihan, C. Poupon, A. Amadon, and F. Lethimonnier, "Artifacts and Pitfalls in Diffusion MRI," *Journal of Magnetic Resonance Imaging*, vol. 24, pp. 478–488, 2006.

[70] O. Pasternak, N. Sochen, Y. Gur, N. Intrator, and Y. Assaf, "Free Water Elimination and Mapping from Diffusion MRI," *Magnetic Resonance in Medicine*, vol. 62, no. 3, pp. 717–730, 2009.

[71] D. Alexander, "Monte-Carlo Studies of Multiple-Fibre Reconstruction Algorithms for Diffusion MRI," in *International Society for Magnetic Resonance in Medicine*, Berkeley, CA, US, Oct. 2005, p. 1344.

[72] W. E. Smith, C.R., Grandy Jr., "Maximum-Entropy and Bayesian Methods in Inverse Problems," in *Acta Applicandae Mathematica*, 1st ed. Dordrecht, the Netherlands: Kluwer Academic Publishers, 1990, pp. 189–191.

[73] T. Schultz, C.-F. Westin, and G. Kindlmann, "Multi-Diffusion-Tensor Fitting via Spherical Deconvolution: A Unifying Framework." in *Medical Image Computing and Computer-Assisted Intervention*, vol. 13, no. Pt 1, Beijing, China, Jan. 2010, pp. 674–681.

[74] C. Xu and J. Prince, "Gradient Vector Flow: A New External Force for Snakes," in *IEEE Computer Society Conference*, San Juan, PR, USA, 1997, pp. 66–71.

[75] W. Freeden and F. Schreiner, "Orthogonal and Non-Orthogonal Multiresolution Analysis, Scale Discrete and Exact Fully Discrete Wavelet Transform on the Sphere," *Constructive Approximation*, vol. 14, pp. 493–515, 1998.

[76] G. D. Parker, D. Marshall, P. L. Rosin, N. Drage, S. Richmond, and D. K. Jones, "A Pitfall in the Reconstruction of Fibre ODFs Using Spherical Deconvolution of Diffusion MRI Data," *NeuroImage*, vol. 65, pp. 433–448, 2013.

[77] O. Michailovich and D. Adam, "A High-Resolution Technique for Ultrasound Harmonic Imaging Using Sparse Representations in Gabor Frames." *IEEE Transactions on Medical Imaging*, vol. 21, no. 12, pp. 1490–1503, Dec. 2002.

[78] J. L. Starck, M. Elad, and D. Donoho, "Redundant Multiscale Transforms and Their Application for Morphological Component Analysis," *Advances in Imaging and Electron Physics*, vol. 132, 2004.

[79] I. Daubechies, M. Defrise, and C. DeMol, "An Iterative Thresholding Algorithm for Linear Inverse Problems with a Sparsity Constraint," *Communications on Pure and Applied Mathematics*, vol. 57, no. 11, pp. 1413–1457, 2004.

[80] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear Total Variation Based Noise Removal Algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1-4, pp. 259–268, Nov. 1992.

[81] S. Boyd, "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, Jan. 2010.

[82] S. Boyd and L. Vandenberghe, *Convex Optimization*, 1st ed. Cambridge, UK: Cambridge University Press, 2004.

[83] A. Chambolle and P.-L. Lions, "Image Recovery via Total Variation Minimization and Related Problems," *Numerische Mathematik*, vol. 76, no. 2, pp. 167–188, 1997.

[84] E. B. Saff and A. B. J. Kuijlaars, "Distributing Many Points on a Sphere," *The Mathematical Intelligencer*, vol. 19, no. 1, pp. 5–11, 1997.

[85] J. P. Haldar and R. M. Leahy, "Linear Transforms for Fourier Data on the Sphere: Application to High Angular Resolution Diffusion MRI of the Brain." *NeuroImage*, vol. 71, pp. 233–247, May 2013.

[86] I. Daubechies, B. Han, A. Ron, and Z. Shen, "Framelets: MRA-Based Constructions of Wavelet Frames," *Applied and Computational Harmonic Analysis*, vol. 14, no. 1, pp. 1–46, 2003.

[87] S. Dolui, I. C. Salgado Patarroyo, O. V. Michailovich, and Y. Rathi, "Reconstruction of HARDI Using Compressed Sensing and Its Application to Contrast HARDI," in *Mathematical Methods in Biomedical Image Analysis*, Breckenridge, CO, USA, Jan. 2012, pp. 17–23.

[88] Y. Rathi, O. Pasternak, P. Savadjiev, O. Michailovich, S. Bouix, M. Kubicki, C.-F. Westin, N. Makris, and M. E. Shenton, "Gray Matter Alterations in Early Aging: A Diffusion Magnetic Resonance Imaging Study." *Human Brain Mapping*, Dec. 2013.